Factors Affecting Automation of Cyber Forensics Investigation

A Dissertation presented to the

Department of Information Systems

University of Cape Town

By

Dean Hayes

(HYSDEA002)

Supervisor: Professor Michael Kyobe

In fulfilment of the requirements for the
Master of Commerce degree in Information Systems

November 2022

DECLARATION

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

2. I have used the APA convention for citation and referencing. Each contribution to, and quotation in, this Literature Review*, The Most Important Factors affecting Automation of Cyber Forensics Phases,* from the work(s) of other people has been attributed and has been cited and referenced.

3. This dissertation, *The Most Important Factors affecting Automation of Cyber Forensics Investigation* is my own work.

4. I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Signature: Signed by candidate ..

Date: 31.10.2022

Full Name of Student: Dean Hayes
Student Number: HYSDEA002

ACKNOWLEDGEMENTS

The global COVID-19 pandemic marks uncertain times and it was not easy to complete a master's dissertation. I humbly give praise to the Almighty Father for giving me strength and courage to achieve this degree.

A huge thank you to my wife Tania Hayes, my mother, my brother and my daughter for their unconditional love and support. I could not have accomplished this degree, without your love and guidance.

My supervisor, Prof Kyobe, there are no words to describe your phenomenal mentorship, kindness, patience, and support. Thank you for holding my hand and pushing me to achieve my goals.

DEDICATION

In loving memory of my dad, Edwin Hayes, and grandma, ‘ouma Henna Kieswetter’ who guides me from the heavens above.

This dissertation is dedicated to my late Father Edwin Hayes and my Grandmother Henna Kieswetter, guiding me from Heaven.

ABSTRACT

Over the past three decades, technology has evolved in a global context changing human interaction through digitization. While there are positive attributes to technological advancement, there are considerable negative elements as well. Cybercrime and digital crime have risen drastically propelling sophisticated digital forensic technology, aimed at fighting cybercrime. However, with automation in digitized Fourth Industrial Revolution, it leaves room to consider the challenges of this on cyber forensic specialists.

The purpose of this study was two folded. Firstly, it aimed to identify factors that affect the automation in the cyber forensics investigation. Secondly, it aimed to determine the most important factors in affecting the automation of cyber forensics investigation performance. The research aimed to shed perspective on automation within the lens of cyber forensics, to provide new insights and efficiency regarding cyber forensics processes. It used a positivism research philosophy and quantitative research design.

Eleven factors that affect the automation of cyber forensics investigation were identified from literature. Furthermore, eleven hypotheses were generated and were tested using a correlation matrix. Additionally, multiple regression analysis was also used to determine the most important factors that affect the automation of cyber forensics investigation performance. In view of this, the R^2 and R^2 change were used to determine to what extent the factors influenced the variance in the automation of cyber forensics investigation performance. In this regard, accessibility to data and dependency on Cloud Service Providers (CSPs) were indicated to be the most important factors that had a significant impact on the automation of cyber forensics investigation performance. It was revealed that accessibility to data influence 37.8% of the variation found in the automation of cyber forensics investigation whilst dependency on CSPs was found to influence 17.5% of the variation found in the automation of cyber forensics investigation. On the other hand, the rest of the nine factors each had a combined influence of 1.7% on the variation found in the automation of cyber forensics investigation performance.

In light of this, accessibility to data and dependency on CSPs were found to be the most important factors that positively affected the automation of cyber forensics investigation performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AI	Artificial Intelligence
API	Application Programming Interface
CAD	Computer Aided Design
CAM	Computer Aided Manufacturing
CSPs	Cloud Service Providers
DFaaS	Digital Forensics as a Service
GDFM	General Digital Forensic Model
IaaS	Infrastructure as a Service
ML	Machine Learning
NIST	National Institute of Standards and Technology
PBF	Push-button forensics
RPA	Robotics Process Automation
SDFIM	Systematic Digital Forensics Investigation Model

CHAPTER 1: INTRODUCTION & BACKGROUND

1.1 Introduction

The birth of cyber forensics is marked in 1976 by Donn Parker's book titled, *Crime by Computer.* The late 1970's and 1980's was an era in which computers were used to process data. Between 1985 to 1995, Hafner and Markoff (1991) indicate that cyber-criminal activities in this time period was young individuals using unauthorised telecommunication networks. Over time the advent of technology advanced, giving rise to cybercrime activities (Pollitt, 2010). In view of extensive cybercrime activities, cyber-forensics rapidly evolved to meet the challenges of the cyber-crime sphere. In the present day, cyber forensics depends on cloud technology and automation, which can also be considered as key contributing factors, to its growth. Yaqoob, Hashem, Ahmed, Kazmi and Hong (2019) define cyber forensics as a science that is related to examination and analysis of electronic data/evidence. This process is composed of four integral phases which are: i) identification phase, ii) collection & preservation, iii) examination & analysis phase and iv) the presentation phase. The chain of custody regulates the process of evidence, from its initial data collection up until presentation. Push-button forensics (PBF) is whereby a forensic investigator uses this forensic tool to achieve the desired results (Horsman, 2020).

Forensics investigation process has modernised with time as technological advancement and currently automation of the process is being explored with the aim of reduce cost and time required for labour intensive manual processes (Hayes and Kyobe, 2020). given that many investigators are being flooded with a lot of cases that they are unable to handle properly resulting in poor quality work outcomes and unwarranted delays (Garfinkel, 2010). The need for automation of cyber forensics investigation process more efficient and could also help to utilise and optimise the existing investigators to gain a better return on training investment made on investigators (James and Gladyshev, 2013). However, automation of cyber forensics investigation process is not an easy endeavour as they are many obstacles to be tackled before it can be done successfully. For instance, the early phases of the investigation process, namely; the data identification and collection phases in many cases than not require human invention and decision making which then makes it difficult to automate these phases (Hayes and Kyobe, 2020).

In this regard investigation phases that need to be automated must be carefully selected by the investigator. According to James and Gladyshev (2013), data preservation, analysis and documentation phases are relatively easier to automate and generate outputs that can be

used in court proceedings. By being able to automate some of the cyber forensics' investigation processes, it may give more time for the investigators to focus on tasks that need to be conducted manually and/or activities that require expert level knowledge to be applied. However, the automation of cyber forensics investigation is not without challenges and these include data complexity, diversity, consistency, volume and unified timelining issues (Boutell and Luo, 2005; Garfinkel, Parker-Wood, Huynh and Migletz, 2010; Raghavan, 2013; Richard III and Roussev, 2006).

Moreover, these challenges may make it difficult to decipher some insights from the data and as a result, some evidence may be overlooked during the automation process particularly when unique scenarios are presented that are beyond the scope of the automation algorithm (Goss and Gladyshev, 2010). Additionally, even though the forensics investigation automation tools may generate useful evidence, its court presentation still requires an expert investigator to authenticate it as presenting the evidence using a non-expert witness can cast doubt on the validity and authenticity of evidence making it to lose its credibility and making it inadmissible in the court of law (Ademu, 2013). Also, automating cyber forensics investigation relies on the existence of electronic information and therefore, organisational readiness for cyber security events also affects the automation of the data as the investigator can only with what the situation presents and the resources as his/her disposal (Lovato, 2017; Sachowski, 2019).

In light of this, the study seeks to present and explain the factors that affect automation cyber forensic investigation and potentially generate new insights in this field of study.

1.2 Rationale of the study

The cyber forensics investigation industry is increasing becoming complex as more technologies are being introduced over time have to a lot of big data that need to be processed by cyber forensics experts. Furthermore, the speed at which new technologies are being introduced are out pacing the rate of adaptability of cyber forensics techniques (Jusas, Birvinskas and Gahramanov, 2017). This presence a scenario in which cyber forensics investigation process is severely affected.

Besides, the ever growing data is only part of the problem but the time required to collect and process the data is also increasing creating another problem (Lillis, Becker, O'Sullivan and Scanlon, 2016). Also, the increasing use of internet has also made it necessary for a lot of law cases to require digital forensics investigation services. In light of this, the digital forensics investigator has to keep him/herself abreast with the technological changes from both the

software and hardware levels in order to successfully investigate and gather the required evidence in a fast and court admissible way and/or state (Alawadhi, Read, Marrington and Franqueira, 2015; Gerda Síochána Inspectorate, 2015). Therefore, the need to automate cyber forensics investigations is paramount and in order to so do successfully there also a need to look into the factors that affect cyber forensics investigation.

In view of this, a number of studies have attempted to look into the automation of cyber forensics investigation in order to improve the efficiency and effectiveness of the cyber forensics investigation process (Ademu, 2013; Al Fahdi, 2016; Alawadhi, 2019; Arshad, Abdullah, Alawida, Alabdulatif, Abiodun and Riaz, 2022; Homem, 2018a; Verma, Gupta and Chang, 2018). But, there is limited evidence of studies that were done in the context of South Africa and/or that have explored the factors that affect cyber forensics investigation process. As such, this study attempts to explore this research gap and hopefully generate important insights that may be valuable to cyber forensics investigation industry in South Africa and the world at large.

1.3 Aim of study

This study aims to first identify the positive and negative factors that influence automation on cyber forensic phases. Secondly, it aims to assess the significance of these factors. In this regard, the study aims to deliver new perspective and insight on the subject matter, which will provide significant contributions for more efficiency in cyber forensic processes.

1.4 Research questions

The research questions of this study are as follows:

- What are the factors that affect the automation of cyber forensics investigation?
- Which of these factors have a greater influence on the automation of cyber forensics investigation performance?

1.5 Research objectives

The research objectives of this study are as follows:

- To identify factors that affect the automation of cyber forensics investigation,
- To determine the extent to which the factors influence the automation of cyber forensics investigation performance.

1.6 Significance of the study

The study explores the factors that affect the automation of cyber forensics investigation. Identifying and knowing these factors is integral to the security of organisations as it enables them to know where potential problems may come from and how they can improve the cyber security of the organisation. According to Homem (2018b), cyber forensics reinforces the existing traditional preventive security mechanisms and also acts as a backup when the traditional preventive security mechanisms fail to detect malicious, stealthy and sophisticated cybercrime events.

However, it should be noted that the cyber forensics investigation process is largely a manual process or at best quasi-automated requiring a highly skilled cyber forensics experts and a lot of time investment (Homem, 2018b). In view of this, automating the cyber forensics investigation has the potential to greatly help organisations to speed up the cyber forensics investigation process which may potentially lead to higher productivity and profitability of these organisations. Considering this, identifying the factors that affect the automation of cyber forensics investigation is integral. This enables organisations to know which areas of cyber forensics to focus on that would deliver a better return on investment in view of the limited resources that organisations may experience.

Zhang (2019) states that automation enables more efficient processes, indicating a large amount of effort invested in automation for cyber forensic processes. However, the challenge is that technology is always changing and improving; therefore, it may not always be possible to keep up with the technological developments. Nevertheless, the foreign process must be constantly reviewed and updated to keep up with the industry standards for cyber forensics process efficiency. For this reason, Yaqoob *et al.* (2019) emphasize the immense benefits of an efficient and optimised process. Contributions in this study area can significantly improve the quality of evidence admissible in a court of law and may also enhance the overall efficiency of cyber forensics investigation process. In this regard, the findings of this study may be of significance to cyber forensics experts, academics, researchers, and other stakeholders who have an interest in this field of study.

1.7 Chapter organisation

The study was divided into five chapters which were outlined as follows:

Chapter 1 outlines the introduction and background of the study. It established the context of study and highlights the issues, challenges, and importance of automating the cyber forensics

investigation process. Furthermore, it gave the study rationale, aim, objectives as well as the problem statement. These help to create a foundation on which the study was built on.

Chapter 2 outlines the literature of the study. This chapter focused on previous study and theory related to the topic under investigation. The main issues that were covered included theories and concepts related to automation of cyber forensics investigation, discussion of the general digital forensics model and factors affecting the automation of cyber forensics investigations.

Chapter 3 outlines the research methodology of the study in which the research philosophy, research design, research instrument, study population, sampling technique, data collection, data analysis, research validity and reliability and ethical considerations were discussed to give a context on how data was collected and analysed.

Chapter 4 outlines the findings from the data analysis and discusses these findings in relation to existing literature. Furthermore, descriptive, and inferential statistics were used to support the research findings.

Chapter 5 gives the conclusions and recommendations of the study. The main focus of this chapter was to draw up conclusions on the research objectives as well as to give recommendations to the study.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter provides the body of knowledge on automation in cyber forensics. It will present literature on the subject area with a clear conceptual framework for integration of the various concepts. It will also explain objectives that are set out before the researcher conducted the study.

2.2 Automation

Automation has drastically changed and reinvented the industries and organisations that has adopted it. However, this concept of automation was introduced in 1946 in the automotive industry to help in speeding the production processes. It is believed that D. S. Harder coined the term “automation” when he was the engineering manager at Ford Motor Company. Nevertheless, the term has now been widely adopted globally by many production and manufacturing industries and beyond (Ahuett-Garza and Kurfess, 2018)

Consequently, the initial concept of automation was based on mechanical and electromechanical control apparatuses – what we now broadly describe as mechanisation. It is also important to note the automation is generally associated with remote controlled devices but in the past decades it seems it is now being more associated with information systems as they are spearheading developments in automation technologies. While the concept of mechanisation has been adopted in automation to relate to basic machine substitution of human labour, automation is usually distinguished separately as the incorporation of machines into a self-governing structure (Groover, 2016). The association with mechanisation has shifted to a greater association with computerisation (Seppelt and Lee, 2019).

Automation is also often implemented beyond manufacturing in association with several structures in which the mechanical, electrical or computerised behaviour is greatly substituted for human action and knowledge (Billings, 2018). Automation is the transition to automated completion, instead of human activity or control of a work cycle, a system, or equipment. Automation does not merely transition human functions to machines, but requires a profound reorganisation of the work cycle, through which the functions of both the person and the computer are described (Endsley, 2018).

Automation can be characterised by the adoption of a technology associated with carrying out a process through controlled commands, supplemented with automatic feedback control for effective implementation of the commands (David, 2017). The subsequent machine can

function beyond human interference. This technology's advancement has become primarily reliant upon the use of information and communication infrastructure technologies and has become relatively more advanced and complex. As a result, these intelligent systems reflect capabilities, skills and efficiency in execution of tasks that surpasses human capacity (Acemoglu and Restrepo, 2018).

The advancement in automation technology has reached a point where a variety of additional innovations has emerged from it and has gained their own popularity and prominence. Robotics is one such technology; it is a specific automation branch in which the automated computer has some anthropomorphic or human features (Momoh, 2017). Besides, automation has gained ground in all industries (Satchell, 2018) and this was partly due to the coming of artificial intelligent (AI) and machine learning (ML) (Rigger, 2019). Nonetheless, one can state that the main feature of automation is to allow processes to occur automatically in the absence of human execution. In most instances, monitoring the process is also automated. Below are a few types of automation we encounter in our current time.

2.2.1 Fixed automation

Fixed automation is typically applied in basic sequencing operations that linear production process in which equipment is aligned in a fixed order assembly operation. However, fixed automation can become more complicated when several activities are combined and organised into one piece of equipment (Das, Roy and Nampi, 2020). In this regard, fixed automation is typically used in applications where there is: i) high initial investment for custom–engineered equipment; ii) high production rates; and/or iii) rigid adherence to the protocol is required (Hilburn, 2017). Moreover, it is important to note that fixed automation is preferrable in products that have a huge demand which can benefit from high volumes and economies of scales to justify huge capital investments.

2.2.2 Programmable automation

Programmable automation exists in systems in which the equipment has the capacity to transform the order in which production operation are conducted to incorporate various product arrangements. In view of this, the various product arrangements for a particular production run are handled by a software programme that follows a series of rules and regulations encoded to it in such a manner that allows the programme to execution of production instructions seamlessly (Tang, 2019). Furthermore, additional new programmes can be arranged and integrated into the equipment to generate new products when needed. Also, it is important to note that flexible automation is preferrable in products that i) require

high investment in general-purpose equipment, ii) have low production rates relative to fixed automation, iii) require flexibility to deal with changes in product configuration and iv) that are most suitable for batch production (Martinova and Martinov, 2018). Consequently, programmable automation is not suitable for continuous production systems as product configurations typically change with each production run and in view of this, batch production is preferrable. According to Kumar (2019), each batch process may require different tools, different machine arrangements and a new set of production instructions to meet product requirements and specifications. Typical examples of programmable automation include but are not limited to industrial robots and numerically controlled machine tools (United States of America Patent No. 9,880,539, 2018).

2.2.3 Flexible automation

Lienenlüke, Gründel, Storms, Herfs, Konigs and Servos (2018) suggest a flexible automated system that creates a range of items without any time lapses between the transition process of products. No manufacturing time is lost in programmable automation extension. For this reason, it can be noted that, the machine can generate a combination of products, rather than generating products in different types of batches. Polak (2019) indicates that flexible automation is preferable in situations where the products require: i) high investment for a custom-engineered system, ii) continuous production of variable mixtures of products, iii) medium production rates and iv) flexibility to deal with product design variations. Coppini and Saliba (2017) emphasize that these characteristics ensure sustainable production without any downtime, which is a defining characteristic of programmable automation between batches. The process of transforming component programming is through scheduling the programmes offline on a computer. Therefore, schedule time for the next production is not disrupted.

2.2.4 Integrated automation

Aamir (2017) explains that integrated automation intends to decrease complexities of independent automated work processes through streamlined communication between the processes that are being automated. As a result, integrated automation plays an integral role of unifying processes within a production system by several technologies such as Computer Aided Manufacturing (CAM), automated material handling and flexible machining systems.

2.2.5 Industrial automation

This type of automation is technology that effectively manages hazardous task for human labour using numerically controlled equipment, CAM and flexible manufacturing systems. In

this regard, industrial automation can help to improve the production efficiency and safety (United States of America Patent No. 10,139,811, 2018).

2.2.6 Computer Aided Manufacturing (CAM)

CAM is the use of computers and machine to automate manufacturing. Moroz (2018) points out that CAM that has many benefits which include but are not limited to improvement of the consistency of material handling, enhancement production output, and improvement product quality. Moreover, designs done with Computer Aided Design (CAD) can also be automatically reproduced or updated using CAM leading to huge time and cost savings (Moroz, 2018).

2.2.7 Robotics Process Automation (RPA)

RPA is used by developers to write codes utilising Application Programming Interfaces (APIs) to automate tasks. Naz and Iraqi (2019) states that RPA interacts with available information technology and communication infrastructure. This general labour-intensive process is easily integrated with an organisation's information technology and communication infrastructure which include websites and applications. RPA is carried by commands executed by bots and these bots according to Wantoo, Bansal and Kushwaha (2017) are used to automate some of the tasks that should be performed by humans.

2.2.8 Cognitive Intelligence

Cognitive Intelligence is the used of software to automate processes which is intensive in knowledge. Amit (2018) cites several advantages, comprising of lower operating costs, increased customer service and many other advantages. This also includes accuracy for complex business processes based on unstructured information (Amit, 2018).

2.3 Principles and theory of automation

2.3.1 Power source

Contemporary automated systems use electricity, as it is the most widely utilised fuel. The actions carried out by automated systems are i) collection, and ii) transition and positioning. The method includes metal shaping, plastic moulding, electrical signal shifting in a communication devices. Momoh (2017) explain that both actions require the use of energy from one state to turn the object which could be a plastic, metal or data into another more desirable state or condition. Precise placement of the product is usually expected at each processing site. Electronic communication is a concept of transferring and positioning in

relation to the transmission of data, between different processing units. It also distributes information to output terminals for analysis and use by human beings.

2.3.2 Feedback controls

Modern automated systems use feedback controls consisting of five basic components. Firstly, *input* to the system is the reference value or set point for the system output. This would be an operating value that should be used to obtain a desired output. For instance, a heating system requires an input for a desired room temperature setting. These feedback controls can also be applicable in a production and manufacturing operation system. Secondly, the *process* is controlled. After processing an input, an output is generated. Once an output is available, *sensing elements* (or measuring devices), are used to measure or monitor the value of an output in the feedback loop. Lastly, *controller and actuating* devices are used to determine whether the input value is producing the desire output value as expected and if not they act to correct or eliminate the difference between them.

2.3.3 Machine programming

Programme directions is a specified series of actions which a device must immediately execute. Suzuki (2017) stresses that the software determines what an automated system can do, and how its various elements achieve an intended outcome. However, it should be noted that machine programming can be quite different from one system to the other. The programme typically consists of limited well-defined actions that perform continuously in sequence without any deviation from one cycle to the other. Furthermore, automation is essential for cybersecurity, to detect, investigate and provide solutions to threats posed by cyber criminality with or without human action. This is achieved by recognising incoming threats, triaging, and prioritising alerts as it emerges. Automation also allows reaction in a timely manner. The advantage of automation is that it can be integrated into cyber security techniques. Automation allows for faster and more efficient data processes. It integrates artificial intelligence into machines, essential for cyber threats. It also eliminates non-cognitive tasks, thereby allowing IT security experts to emphasize on more important aspects.

2.4 The concept of cyber forensics

It can be noted that historical developments of cyber forensics influenced future developments of cyber forensics as a specialise field. Cyber forensics umbrellas digital forensics, detecting, extracting, examining and validating information regarding digital evidence (Karafili, Cristani and Viganò, 2018).

The historical development of cyber forensics heavily influenced the direction it took in future developments, resulting in the formation of cyber forensics as a specialised field of its own, and established as an academic discipline. Cyber forensics can also be defined as digital forensics - a category of forensics science that involves detecting, extracting, examining, validating and presenting information about digital evidence stored on computers or related digital media storage devices (Karafili *et al.*, 2018).

Cyber forensics uses electronic data for detection and analysis. The procedure captures and retains evidence in its initial form through a systematic inquiry of documentation, reinforced by digital information to recreate actual events (Husain and Khan, 2019). Cyber forensics also preserves, collects, validates and interprets digital evidence obtained by digital sources. It aims to facilitate reconstruction of events essential to address cybercrimes (Husain and Khan, 2019).

Cyber forensics ensures the forensics team has the most suitable techniques and tools to address complex digital related cases. It is important to note that definitions of cyber forensics vary across different applications. However, the common thread is that computer forensics, digital forensics or cyber forensics are mainly adopted within the literature as concepts that relate to the sub-branch of forensics science (Shrivastava, Sharma, Khari and Zohora, 2018). Since cyber forensics forms part of forensics science, it is mainly concerned with features of legal evidence that can be obtained from utilising computers and other digital related devices. In some instances, it is also referred to as computer forensics and could consist of sub-branches that include network forensics and mobile forensics (Glisson and Choo, 2018). It is a method for obtaining evidence, retaining data lost, establishing how the security breaches were executed and setting a sequence or timeline to digital events. Therefore, it is regarded as a mechanism to obtains evidence on computer related crimes. For instance, hacking, fraud and identity theft.

Digital forensics in an investigation that is seen as a series of tasks and processes (Pollitt, 2010). However, the process list is very restricted and reconstruction is required at times. Reconstruction is classified as an item that locates a root cause to simulate projected events that leads to a state (Yadav, 2020). It is based on objective study that supports scientific principles, and that should be replicable.

Significance is also given to the tools, as opposed to solely relying on methods that are adopted to generate the evidence in cyber forensic investigations. Whilst some definitions merely focus on the criminal element, the scope is broadened to classify cyber forensics as a

mechanism for attaining initial causes in other categories of investigations (Chhabra, Singh and Singh, 2020). However, the main causes are not a part of the definition. The basis for this is that specific investigations might not have particular root causes (Arnanda, 2020).

One can regard cyber forensics based on mechanisms and instruments of scientific principles, applicable to specific digital media or digital data (in a particular predefined process or accepted process). It also takes into consideration legal principles. It places emphasis on obtaining digital evidence and specifying a series of events or actions as the main causes in order to support the reconstruction, decision and future prevention of cyber misuse (Al-khateeb and Agarwal, 2019).

2.5 Different models/frameworks of cyber forensics investigation

There is a general misconception of no universal conceptual approach for the application of digital forensics. This is based on the fact that each case has novel features and requirements that cannot necessarily be applied elsewhere since specific benchmarks and standards are meaningless (Taylor, Fritsch, Liederbach, Saylor and Tafoya, 2019). With reference to the fast shift of technology, digital forensics models have to be based on vital principles and processes that will be essential for the future, regardless of techniques and mechanisms frequently shifting (Tonye, 2018). Through focusing on the model, principles and concepts as opposed to detailed lists and procedures, the frameworks can be applicable to various environments and situations.

In 1999, McKemmish proposed a digital forensics framework that has four phases which are i) identification phase, ii) preservation phase, iii) analysis phase and iv) evidence presentation phase (Du, Le-Khac and Scanlon, 2017). However, Du *et al.* (2017) advanced the event-based forensics investigation framework, which consists of preservation, search and reconstruction phases. The preservation phase preserves the state of the system, where primitive and complex capability definition techniques would be used. Beebe and Clark in 2004 advance the hierarchical, Objectives-Based Framework which consist of two hierarchical tiers, with the first tier being more focussed on physical implementations (Du *et al.*, 2017). They propose this to allow lower-level objectives and processes to be presented as part of the model. One can argue that this is justifiable due to the extent of some criticism of prior models that are regarded too broad and being detached from the practical universe. Such models have emphasised vital ideas that can be adopted as guidelines for academics and practitioners to be used within novel circumstances. It is vital to note, there is continuous discussions regarding the model which can be used within digital forensics (Jaishankar, 2018).

2.5.1 Improved digital investigation process

Baryamureeba and Tushabe (2004) proposed a change to Integrated Digital Investigation Model developed by Carrier and Spafford who cited two additional stages: trace back and dynamite. These additional stages have the intention to differentiate the investigation of the computer and of the physical scene. The main purpose of this differentiation will be to help the investigators to reconstruct computer and physical crime scenes in such a way that eliminates inconsistencies. For this reason, this theory or process was not suitable.

2.5.2 Extended model of cybercrime investigation

Kao (2016) asserts that current models are universal models of cybercrime investigation that focus only on processing of evidence in cybercrime investigation. The model offers a good foundation to understand the investigation process and the undertaking of the information flows (Palmer, Llorens, Kaufman, Gibbons, Chowdhury, Chen and Fu, 2016). Nevertheless, the model is generic and focuses on the management aspect and thus it was not suitable.

2.5.3 Case-relevance information investigation

Rughani (2017) argues that there is a need for computer intelligence technology to present computer forensics frameworks. They put forward that computer intelligence is expected to assist in the investigation processes by providing and sharing knowledge efficiency with other investigators during an investigation or across various cases (Kothari and Hasija, 2017). The initial concept that was introduced by the authors is the notion of attaining knowledge, that is the investigative benchmark that determines analysis of data. An additional concept described by the authors is the concept of Case-Relevance. They adopted this aspect to define the difference between computer security and forensics, regardless of defining degrees of case relevance.

2.5.4 The Systematic Digital Forensics Investigation model (SDFIM)

Agarwal, Gupta, Gupta and Gupta (2011) advanced a model with a purpose to assist forensic practitioners and organisations for structuring suitable policies and procedures in an organised manner. The proposed model emphasises an investigation on computer fraud and cybercrimes. Applying the concepts and suggestions of this model is mainly confined to computer fraud and cybercrimes. The digital forensic investigation will be separated into a 4-tier iterative processes which consist of preparation tier, collection tier, examination tier and presentation tier (Koleoso, 2018).

2.5.5 McKemmish 1999 Model

McKemmish (1999) defines cyber forensics as the process of recognising, preserving, examining and presenting digital evidence in a way that is legally recognisable. This involves four phases which are i) identification, ii) preservation, iii) analysis and iv) presentation of digital evidence (Lutui, 2016). The identification phase involves gathering information about the evidence about its availability, its location, its categorisation and its format. On the other hand, the preservation phase focuses on ensuring that the integrity of evidence and that it is not altered, changed or modified in any manner during the investigation process. In this regard, a cyber forensics investigator should not tamper with evidence during his or her investigation process (Mody and Nisbet, 2017). In addition, the analysis stage changes the bit level data obtained in the earlier two phases into evidence which can be used in court of law. The final phase, which is presentation phase focuses on presenting the evidence to courts of justice in accordance with presenting expert testimony on the analysis of the evidence. However, this model deems to be outdated and for this reason, it is not considered.

2.5.6 The National Institute of Standards and Technology (NIST)

According to Conlan, Baggili and Breitinger (2016) another broadly adopted digital forensics framework is the one by NIST. The four phases and definitions within the NIST framework have identical features with the one provided by McKemmish (1999). These phases include collecting and recognising essential data, preserving its authenticity and obtaining the data (Choo, 2017). Examination data adopts automated and manual tools to extract data from the internet, whilst ensuring preservation. Analysis is emphasised on deriving essential information from the results of the examination, the last phase, reporting is focused on preparing and presenting the forensic analysis.

2.6 The General Digital Forensic Model (GDFM)

Technological advancement may change or alter tools and methods that are employed in cyber forensics investigation. This is necessary because cyber forensics investigation processes must remain relevant and up to date with the current technologies that may be used against the organisation if it is not adequately prepared to combat them (Rogers and Seigfried, 2004). In view of this, the study uses the GDFM as it captures this concept and helps the cyber forensics investigator to identify and rank the factors that affect the automation of cyber forensics processes as it describes the cyber forensics investigation process and it gives consideration to the forensics client, forensic element and forensic processes that are all integral in the investigation process. The GDFM is shown in Figure 1 below:

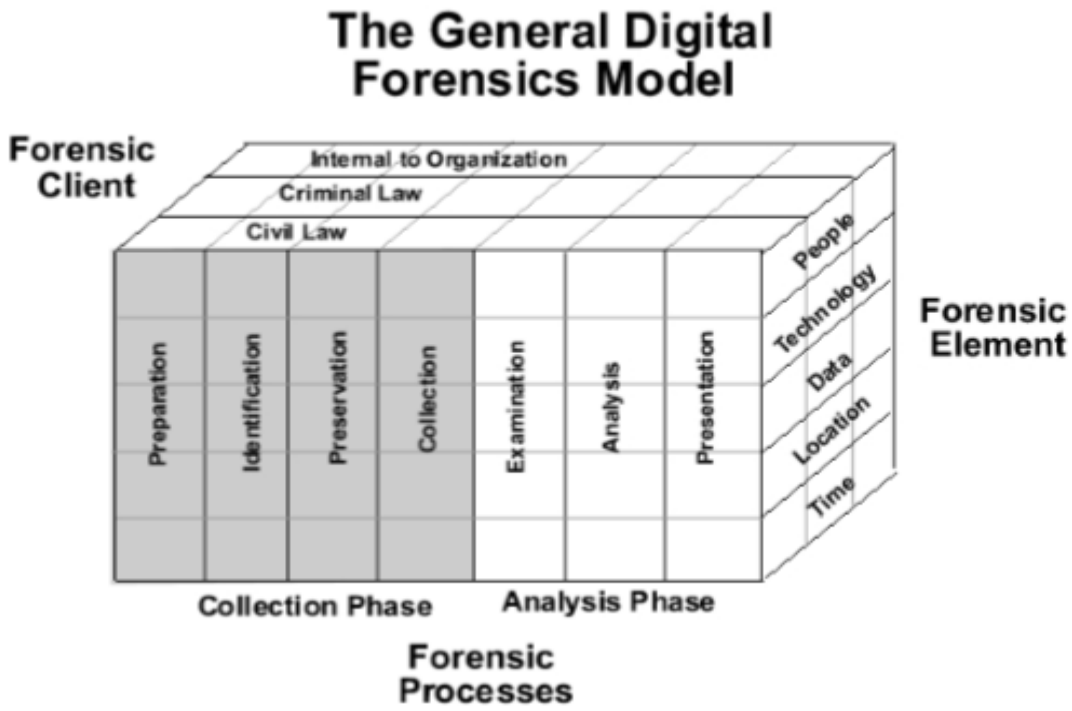

Figure 1: The General Digital Forensics Model

Source: Rigby and Rogers (2007)

2.6.1 Forensic elements

The forensic elements as shown by the GDFM in Figure 1 consist of people, technology, data, location, and time. These five forensic elements are integral in ensuring the efficacy of the automation cyber forensics investigation process. Firstly, the people element is related to individuals who are involved in the cyber forensics investigation processes and these may include cyber forensics investigators and lawyers which typically perform and guide the cyber forensics investigation process. In view of this, automation may alleviate some of the tasks from the people involved and this also transfers the responsibility and accountability related those tasks to the automation system (Al-khateeb and Agarwal, 2019). Secondly, the technology element is concerned with the software and hardware that the cyber forensics investigator would use to perform the investigation process which include but not limited to hardware write-blockers. Generally, it is believed that technology becomes obsolete within a 2 year timeline (De Mauro, Greco and Grimaldi, 2016). This suggests that forensics tools with new capabilities should be developed ahead of the technological advancement to curb cybercrimes before they occurred. Though this is a desirable wish, it is sadly not the case and cyber forensics investigators usually have to fight cybercrime from a position of technological disadvantage. As this is not enough, the heterogeneity of the global IT infrastructure does not

help the automation of cyber forensics investigation (Lu and Da Xu, 2018). Also, it is important to always check and validate the accuracy and efficacy of cyber forensics tools to ensure that they are operating as intended (McCartney & Amoako 2018). This validation process should be done by an expert and operational step may be done by a non-expert user, but a challenge may be experienced when the non-expert user is called to testify in the court but does not have full knowledge about how the evidence is collected. This can create credibility issues in the court of law which may affect the weight given to the evidence provided (Horsman, 2020).

Thirdly, the data element relates to different categories that the cyber forensics investigator may encounter during the investigative process. Consequently, during the identification phase the influence of technology on the automation of the cyber forensics' investigation process is not that strong. However, Digital Forensics as a Service (DFaaS) can facilitate the gathering of evidence without tampering with it (Bhandayker, 2019). Since data integrity must be preserved, it is necessary to employ cryptographic hash to the data to ensure that it is not changed, altered or modified. Furthermore, if physical location of the evidence is unknown then challenges in identifying the evidence are to be expected. In this regard, if the location of the data is unknown it may become impossible to automate the process with smart algorithms. However, if the location of the evidence is known and it is identified on hard drives, cloud infrastructure and software applications smart algorithms can be used to automate the data collection and preservation phases whilst simultaneously maintaining the chain of custody of data (Arafat, Mondal and Rani, 2017). This is very important because if the chain of custody of the evidence is not maintained then the data would be acceptable in the court of law. It is also important to note that automation tools should have the ability to search, collect and analyse big data without being limited by storage space (Horsman, 2020).

Additionally, it should be noted that it is important to maintain the privacy during the collection and preservation phase as smart algorithms should not be able to collect confidential data without consent as infringes the privacy laws and discredits the evidence whilst also exposing the cyber forensics investigator to lawsuits (Horsman, 2020). Furthermore, it should be noted that it is not possible to automate the whole cyber forensics investigation without missing some evidence and the cyber forensics investigator should be cognisance of this and should act in a manner that complements the automation tools (Horsman, 2020). Fourthly, the location element deals with identification of where the cybercrime occurred, and this may affect the investigation process as different countries may require the cyber forensics investigation process to be conducted to conform to the legislation. Lastly, the time element deals with the length of time required to conduct the cyber forensics investigation process from start to finish. According to (Govindaraj, Verma & Gupta 2018), automation of security analysis and using

forensically ready solutions which are able to learn and adapt to new forensics challenges can help to improve the efficiency and productivity of the cyber forensics investigation.

2.6.2 Forensics processes

The GDFM portrayed in Figure 1 shows that the forensics processes is composed of two sequential phases which are the collection phase and analysis phase. Accordingly, the collection phase has four stages which are preparation, identification, preservation and collection. Whilst the analysis phase has three stages which are examination, analysis and presentation of the evidence in a justice court.

2.6.3 Forensics client

The forensic client is composed of three items which are internal organisation, criminal law and civil law as shown in the GDFM in Figure 1. It is important to distinguish whether an investigation is civil or criminal. Criminal investigation will land up in court of law whereas civil cases do not. The internal organisation can be referred to as the actual company or employees the cybercrime has affected or even if internal staff are involved in the crime.

Unlike other models, the GDFM describes the inter-relationships between the forensics process, forensic elements, and forensics client. The GDFM relates directly to the cyber forensics investigator by identifying processes and principles which consists of the collection phase and analysis phase. The relationship of the forensic elements is also highlighted. It is also evident that the forensic client concept has relationships with both; forensics processes and forensic elements. The forensic client consists of the different law factors and the internal organisational factors that cuts over the forensic processes and elements (Rigby & Rogers, 2007). This research will cross over all these concepts, as the researcher will identify the factors that affect the automation of the cyber forensics process.

While the Technology, Organisation and Environment (TOE) model was initially considered to highlight the factors that could affect the automation of the cyber forensics process, the model lacks the people aspect, as forensics investigators and lawyers play a vital part in the process. The TOE model also does not relate the theory to the various forensics phases and the interlinking of it with the elements, thus the GDFM was chosen as the best representation for the present research study. The GDFM provides a more comprehensive view of the cyber forensics process, making up for where the TOE is lacking. The GDFM is multi-dimensional and graphically highlights the cyber forensics process and the factors that influence it. It also

stimulates careful considerations that helps students and professionals seriously rethink the cyber forensics process and the inter-relations of forensics elements and forensics clients.

2.6.4 Forensic client (forensic law elements)

The forensic client can also be referred to the forensic law element that the role of forensics investigators, lawyers and internal organisation aligned to the type of law associated with the cyber forensics process and the automation of it (Du, Le-Khac, & Scanlon, 2017). These categories of law are known as criminal and civil lawsuits, which will be unpacked later in this chapter. Organisational factors also have an impact during the forensic investigation.

2.6.4.1 Internal Organisation

One of the forensic clients or forensic law elements that is mentioned on the model speaks to organisational factors such as i) managerial structure, size or culture of the firm, ii) communication, iii) external factors (legal challenges, competitors and other external stakeholders) in adopting automation in the cyber forensics process. Literature review did not yield a definition, with clear objectives and methodologies for adopting automation on the cyber forensics phases for the organisational factors (Du *et al.*, 2017).

The ownership of cyber security has moved from being led and operated by IT within a company to board of directors' ownership. The board of directors is involved and they direct and support these initiatives. The risk has moved from an operational impact to more of a reputational compliance and financial impact. The investment for a company is moved from being technical, to a protective culture and threats have moved from being mainly external to both external and internal. One of the most common denominators of organisations these days are to stop hackers and to encourage companies to look inside their walls. Many of these breaches are due to insider human error. Organisations invest in traditional security like hardware, firewalls, and software protection in email filtering etc., but the biggest risk and the biggest line of defence are people and it's what we call the human firewall.

From an external threat, the range of threat actors are diverse, and the control and threat mitigation landscape are ever-growing. Knowing what and who to combat using different tools and methods can be very challenging for the most part, as these external threat actors are hacktivists, criminals and corporate espionage, and for the most part they usually share one common goal; to disrupt data or acquire that of value. Although motivations are somewhat different from an external perspective, they usually have the same hurdles to overcome. First

a weakness in the system needs to be found to get a foothold on the environment. Organisational structure greatly dictates how easy it is for external threats to have an impact.

It can be said that intellectual property theft and insider theft can influence internal factors of an organisation when applying automation to the cyber forensics' phases. Placing trust in employees is an important organisational cultural factor in not only attracting but also retaining talented individuals and for encouraging a positive and collaborative corporate culture. At the same time, it serves as an internal threat. Management must verify what employees are doing before they potentially walk out the door with the company's intellectual property – civil litigation can be a costly but avoidable process. President Ronald Reagan, while discussing US relations with the Soviet Union, would famously quote the Russian proverb 'trust but verify' and when it comes to retaining valuable intellectual property for your company, it might be useful to use that same proverb as a mantra in the future.

The availability of forensics investigators and lawyers to work in uniform, particularly in economic struggling countries are lacking. Ultimately this will have a negative impact for the cyber forensics' profession and decrease the chances of automation of the cyber forensics process. Cyber criminals take advantage of this, knowing the limited resources to combat cybercrime in the economic struggling countries of the world and in turn initiating numerous crucial cyber-attacks (Bhandayker, 2019). The forensic investigators and lawyers should support each other and collaborate so that the forensics process can be standardised and automated for better use. Cyber forensics has a significant place in criminal and civil investigations.

2.6.4.2 Civil Law

Digital information has become an essential element in civil cases. Across the world there are countless forms of electronically stored data but often an entire legal case can hinge upon a single record such as a text message, voicemail, spreadsheet, or database containing trade secret information. A generation ago evidence in civil cases was commonly found in overstuffed filing cabinets, day planners and locked office drawers. The rise of cyber forensics coincides with the change in global business practices. Computers, mobile phones, servers, and cloud storage are now the data repository's housing sensitive data. Electronic discovery (E-Discovery) has specific legal limitations and restrictions usually in relation to the scope of any investigation. Privacy laws for example the right of employees not to have personal conversation intercepted and human rights legislation often affect electronic discovery.

The first step in the process begins with a thorough consultation so that it can determine how cyber forensics expertise can be of most use in this phase. Closely work with the information technology professionals at the organisation to determine the types and location of electronic media that have potential interest. Next comes the acquisition process where electronic data, using forensics tools and methodologies, gets collected in accordance with cyber forensics best practices ensuring that the integrity of the data is maintained and will hold up in court. Data is collected from various sources using the best methods possible.

Once the data is collected, we move to the analysis phase using specialised forensics technology and methods. Our experts examine the data including the recovery of deleted data. During the in-depth analysis process, it accurately determines what and how it occurred and who the responsible parties are. At the completion of the analysis, reporting begins at the client's request. Technical details can be complex and that is why comprehensive reporting that is well written and understandable, translating difficult concepts into plain language is imperative. Finally, we have expert testimony to explain forensics procedures, digital artifacts, and technical concepts in a compelling and accessible way to the court or client.

Civil law follows different standards to that of criminal law, namely preponderance of the evidence. If there is a 51% certainty of negligence, then the individual is liable. Thus, the evidence standard is much easier and what's at stake is money or collecting monetary damage. Criminal cases generally have a higher burden of proof than civil cases, this is mainly because a person's freedom is at stake.

2.6.4.3 Criminal Law

There is a distinction between civil and criminal law. The terminology 'criminal' is related to guilt and 'civil' is related to being liable, and consequently the way evidence is treated in these cases also differ. In criminal cases guilt is found when evidence is clear beyond a reasonable doubt. Cyber forensics is traditionally associated with criminal investigations, as would be expected. Most types of investigation encompass some form of computer crime. This sort of crime can take 2 forms; computer-based crime and computer facilitated crime. Computer based crime is criminal activity that is conducted purely on computers for example cyber bullying or spam as well as crimes newly defined by the computing age. It also includes traditional crime conducted purely on computers for example child pornography.

Computer facilitated crime is conducted in the real world but facilitated using computers, a classic example of this sort of crime is fraud. Computers are commonly used to communicate with other fraudsters to record planned activities or to create fraudulent documents and not all

cyber forensics investigations focus on criminal behaviour. The techniques that are used in corporate or private settings is to recover lost information or to rebuild the activities of employees. Criminal forensics is the largest form of cyber forensics and falling under the responsibility of law enforcement or private contractors working for them.

Criminal forensics is usually part of a wider investigation conducted by law enforcement and other specialists with reports being intended to facilitate the investigation and ultimately to be passed as expert evidence before the court, which focus on forensically sound data extraction and producing report evidence or in simple terms, intelligence gathering. This type of investigation is often associated with crime but in relation to providing intelligence to help track, stop, or identify criminal activity. Unless the evidence is to be used in court at a later stage, forensic soundness is less of a concern in this form of investigation, instead speed can be a common requirement.

The rigorous requirements for the quality of evidence makes it difficult to be admissible in a court of law. Literature states that the technologies to be able to fully automate complicated evidential analysis and reasoning tasks without human intervention are not yet available, and courts admitting automatically extracted evidence without some sort of human expert verification appears unlikely (Horsman, 2020). Automation of the different cyber forensics' phases might speed up the validation process, but the validity of the automated procedure should be proven (Horsman, 2020).

2.7 Factors influencing cyber forensic automation.

This chapter discusses the factors affecting automation in cyber forensic phases. In view of this, four phases are discussed in the sub sections below, and these are: i) preparation and identification phase, ii) collection and preparation phase, iii) examination and analysis phase and iv) presentation phase as illustrated in Figure 1 above. However, these are discussed in detail in the following subsections below.

2.7.1 Preparation and identification phase

The first phase in the GDFM shown in Figure 1 is to prepare and identify whether there will be an investigation (Rigby and Rogers, 2007). Moreover, it is important to note that not every incident requires an investigation because some of the incidents may occur as a result of a mistake or a natural disaster.

2.7.1.1 Accessibility to data

It can be noted that if data is not filed autonomously, it would mean that investigators would not know if an investigation is needed or not which can be an issue in cases where data is stored in clouds which are distributed over several locations globally (Morales-Ferreira, Santiago-Duran, Gaytan-Diaz, Gonzalez-Compean, Sosa-Sosa and Lopez-Arevalo, 2018). In light of this, human intervention will be necessary to retrieve the data which can be done with the assistance of tools such as Syslog and Log Analyzer which were designed to correlate and maintain logs within the log management system (Rane and Dixit, 2019). This means that Cloud Service Providers (CSPs) should prove logging instruments that help to automate cyber forensics investigation process.

H_{1a}: *Accessibility to data facilitates automation of cyber forensics investigation during the preparation and identification phase.*

2.7.1.2 Unstable data

Normally when a device is turned off, cyber forensics investigators will not be able to automatically identify evidence as the registry and temporary internet files will be lost (MacDermott, Baker and Shi, 2018). Consequently, it will require human intervention to turn the device on either remotely or physically. Although, Damshenas, Dehghantanha, Mahmoud and Bin Shamsuddin (2012) suggested the use of data-recovery and data-safety mechanisms during the identification phase, proven and suitable mechanisms should be used to minimise instances where the integrity of the evidence is ruined.

$H_{1b:}$ *Unstable data impedes the automation of cyber forensics investigation during the preparation and identification phase.*

2.7.2 Collection & preservation phase

The phase of collection and preservation is very important as the collected evidence should be preserved in its original state and measures should be put in place avoid tampering with the evidence and as such verified and safe data collection and preservation mechanisms (Lone and Mir, 2019).

2.7.2.1 Dependency on CSPs

CSPs due to privacy and confidential laws they are not generally willing to share clients information and this makes the automation of cyber forensics investigation nearly impossible. Besides, CSPs also only keep data on their cloud systems for short periods of time and

because of the way their clouds systems are design, data can easily be deleted or overwritten with new data. Additionally, data may also be spread in several data centres globally and CSPs do not display the locations (Morales-Ferreira *et al.*, 2018). Also, when there is a breach on the system, CSPs generally prioritise the restoration of the service rather preservation of evidence for the cyber forensics investigation. Moreover, CSPs also do not feel obligated to report cybercrime incidents as it may reflect bad on their reputation and thus prioritise restoring their cloud systems back online above all things. Essentially, the cloud infrastructure of CSPs is designed for operational use and without cyber forensics investigation in mind. However, different automation tools and models have underwent experimentation in cloud environments to safeguard forensically sound evidence to be collected and preserved. Nonetheless, in order to minimise these challenges Trust Cloud has been developed as Infrastructure as a Service (IaaS) to assist cyber forensics investigators to automatically acquire the evidence they need (Alqahtany, 2017).

H_{1c}: *The use of Trust Cloud facilitates the automation of cyber forensics investigation during the collection and preservation phase.*

2.7.2.2 Minimize time, maximize coverage

In December 2010, DFaaS was introduced by the Netherlands Forensic Institute and has undergone significant advancements over the years (Du *et al.*, 2017). DFaaS is a cloud-based service that allows forensic copies to be made of devices and stored to a central storage that is used to extract or collect evidence automatically for analysis, making a significant contribution to automation of the collection phase. Forensic copies are made from various devices and platforms; thus, a huge area of data is covered. The accessibility of forensic copies ensures that evidence can be collected when the need arises, saving time and resources (Du *et al.*, 2017).

H_{1d}: *DFaaS facilitates the automation of cyber forensics investigation during the collection and preservation phase.*

2.7.3 Examination & analysis phase

The examination and analysis phase involves combing through big data utilising both manual and/or automated mechanisms to extract relevant data which will be analysed using legal techniques to generate important insights that can be presented in a court of law (Yaqoob *et al.*, 2019).

2.7.3.1 Decline in expert knowledge

Examination and analysis in cyber forensics do not typically make use of full automation but does involve a certain level of automation that is used in 'push-button forensics' (PBF). Well known PBF suites are the likes of EnCase, Forensic Tool Kit, BelkaSoft, and Autopsy Forensic Browser. Some investigators make their comfort levels clear in regard to PBF, as it will have a stagnation effect on their knowledge learning in this profession (Horsman, 2020). However, some analysis tasks would be unrealistic without using PBF tools. These PBF tools allow forensic investigators to perform complicated analysis functions, purely by knowing which buttons to press. The higher level of automation PBF tools can influence the quality of the analysis on the evidence, including exposing the lack of skill of the investigator, which in turn can have a negative impact in a court of law. In contrast, conducting the analysis phase manually is reliant on expert skill, and will require the investigators to invest much more of their time and effort (Horsman, 2020).

Automation and arterial intelligence may slow down cyber forensics experts, as they solely rely on automated tools for all data interpretation and analysis tasks, and derived evidence is admitted in court without challenge (Vallor, 2017). The upskilling of forensic investigators and lawyers is an important factor when considering automation in cyber forensics. Certification, continuous cyber forensics laboratory experiments and research is key for the development of a future autonomous cyber forensics process (Garfinkel, 2010). Hitchcock, Le-Khac and Scanlon (2016) suggests specific courses at tertiary institutions to support the ever-growing industry. Continuous training of staff in the forensic department is imperative so that their knowledge is always up to date with cutting-edge forensic tools.

H_{1e}: *PBF facilitates the automation of cyber forensics investigation during the examination and analysis phase.*

2.7.3.2 Hardware requirements

The powerful digital forensics suites are demanding when considering the hardware requirements in terms of memory size, central processing power and disk space. The standard off the shelf computer is not capable of handling this type of churning (Atlam, Alenezi, Alassafi, Alshdadi and Wills, 2020). High performance machines need to be used for automating the analysis phase as it syphers through massive datasets (Irons and Lallie, 2014). In addition to high performance computers, special equipment is also required (Homem, 2016). This suggests that hardware specifications and the quality of evidence and investigation results are interlinked.

H_{1f}: *Hardware requirements impedes the automation of cyber forensics investigation during the examination and analysis phase.*

2.7.3.3 Profiling & event reconstruction

When the cyber forensics process becomes fully automated, including eluding bias and preconception, case backlog will be minimised. The automation process should be validated and evaluated by doing thorough and complete analysis, like profiling (Al Mutawa, Bryce, Franqueira, Marrington and Read, 2019) or reconstructing the incident automatically (Horsman, 2020). Despite a variety of practical techniques and tools, digital forensics provides a small theoretical basis to support and correlate findings of an investigation. In future, when the forensics community grasps formal techniques, tedious ad-hoc validation will continue. Furthermore, the constructive value of the tools and analysis will endure (Horsman, 2020).

H_{1g}: *Profiling and event reconstruction facilitates the automation of cyber forensics investigation during the examination and analysis phase*.

2.7.3.4 AI

Intelligent forensics is an interdisciplinary approach utilising modern technology, contributing to an effective strategy to unravel an examination. As per custom, cyber forensics uses queries to find information. In addition to queries, intelligent forensics uses techniques that enable data to find queries, data to find data and queries to find queries (Wylot, Hauswirth, Cudré-Mauroux and Sakr, 2018). Intelligent forensics consists of a variety of tools and techniques from AI and machine learning. It initiates modelling and social network analysis to focus on digital investigations and cut back on the quantity of the time spent looking for digital proof.

H_{1h}: *AI facilitates the automation of cyber forensics investigation during the examination and analysis phase*.

2.7.4 Presentation phase

This phase is where cyber forensics investigators reap the fruits of their hard work by displaying evidence which will be validated in a court of law (Chhabra *et al.*, 2020).

2.7.4.1 Non-expert investigator

When a non-expert investigator presents evidence that was automatically discovered, it can raise a concern for lawyers and judges. It is important that an expert witness validates all the

forensics software and tools that have been used during the investigation. If this fails to happen, evidence will lose credibility and become inadmissible in court.

H_{1i}: *Non-expert investigator impedes the automation of cyber forensics investigation during the presentation phase.*

2.7.4.2 Spread of data in the cloud

Due to the vast infrastructure of the cloud, expert witnesses find it challenging to prove in which country an incident has occurred. This in turn, creates misalignment with the investigators by establishing which laws to obey based on the specific country. Therefore, investigators must explain the technicalities of how this phase was automated to acquire the evidence from the cloud environment. In turn, the jury find it hard to comprehend complexities of cloud forensics (Simou, Kalloniatis, Gritzalis and Mouratidis, 2016).

H_{1j}: *Spread of data in the cloud impedes the automation of cyber forensics investigation during the presentation phase.*

2.7.4.3 Reliability and privacy

During this phase of automation, reliability and privacy is crucial for the law. The concept of reliability is connected to the accuracy of the forensic tool (Hughes and Karabiyik, 2020). The safe keeping and security of data will always be vulnerable. However, to mitigate this, tools need to drastically improve, and the cyber forensic process should become less cumbersome and more automated (Hughes and Karabiyik, 2020).

H_{1k}: *Reliability and privacy concerns impedes the automation of cyber forensics investigation during the presentation phase.*

2.8 Conceptual framework

The conceptual framework of the study is shown in Figure 2, and it was created from the hypotheses generated in the previous section. The relationship between these factors was tested in this study.

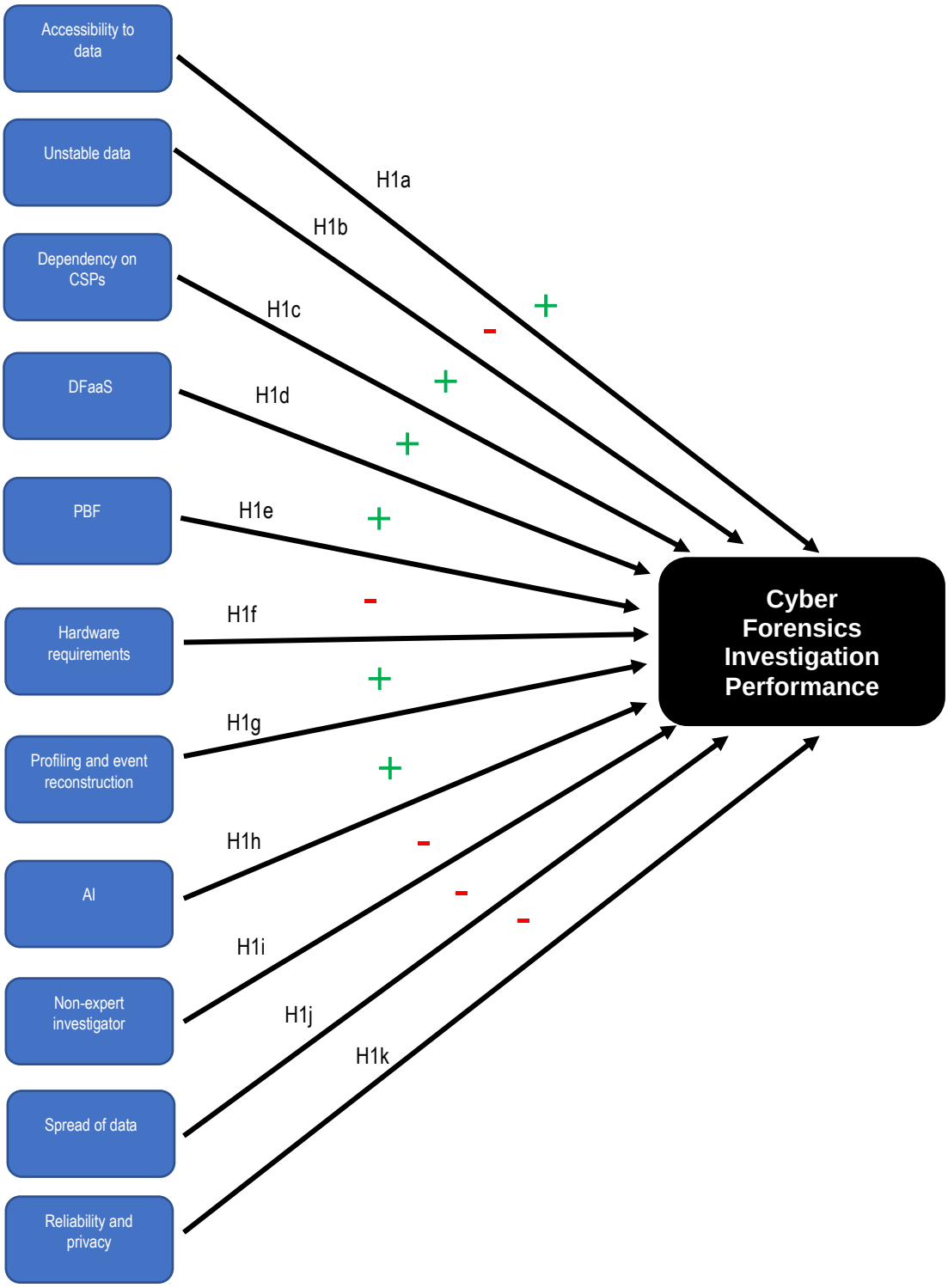

Figure 2: The conceptual framework of cyber forensics investigation performance

2.9 Chapter summary

This chapter discussed the concepts and theories related to the automation of cyber forensics investigation. Furthermore, the general digital forensic model (GDFM) was used to outline the phases involved in the cyber forensics' investigation process. Also, eleven factors that affect cyber forensics investigation processes were identified and these were: 1) Accessibility to data, 2) Unstable data, 3) Dependency on Cloud Service Providers (CSPs), 4) Minimise time, maximise coverage, 5) Decline in expert knowledge, 6) Hardware requirements, 7) Profiling and event construction, 8) Artificial intelligence, 9) Non-expert investigator, 10) Spread of data in the cloud and 11) Reliability and privacy.

Furthermore, the following hypotheses were also generated in the literature review:

- H_{1a}: *Accessibility to data facilitates automation of cyber forensics investigation performance during the preparation and identification phase.*
- $H_{1b:}$ *Unstable data impedes the automation of cyber forensics investigation performance during the preparation and identification phase.*
- H_{1c}: *The use of Trust Cloud facilitates the automation of cyber forensics investigation performance during the collection and preservation phase.*
- H_{1d}: *DFaaS facilitates the automation of cyber forensics investigation performance during the collection and preservation phase.*
- H_{1e}: *PBF facilitates the automation of cyber forensics investigation performance during the examination and analysis phase.*
- H_{1f}: *Hardware requirements impedes the automation of cyber forensics investigation performance during the examination and analysis phase.*
- H_{1g}: *Profiling and event reconstruction facilitates the automation of cyber forensics investigation performance during the examination and analysis phase*.
- H_{1h}: *AI facilitates the automation of cyber forensics investigation performance during the examination and analysis phase*.
- H_{1i}: *Non-expert investigator impedes the automation of cyber forensics investigation performance during the presentation phase.*
- H_{1j}: *Spread of data in the cloud impedes the automation of cyber forensics investigation performance during the presentation phase.*
- H_{1k}: *Reliability and privacy concerns impedes the automation of cyber forensics investigation performance during the presentation phase.*

The next chapter will discuss the research methodology of the study.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

The study followed a cross-sectional, mixed-method design, consisting of a review as well as a survey component that utilises both quantitative and qualitative methods. Literature states that the 'what' and 'how' is descriptive and the 'why' is explanatory (Whetten, 1989). The nature of the research is primarily descriptive and exploratory as the sparse literature does not provide a full view of the cyber forensics process as it has been presented here. It identifies and explores factors highlighted in the literature and furthers the understanding of what the most important factors and conditions are when applying automation to the cyber forensics process by collating expert feedback and rankings of the factors deemed most important. In the next section, the researcher provides more detail of the philosophy, approach, purpose, sampling, data collection and validity of this paper.

3.2 Research philosophy

The philosophy of this study is how the researcher views the cyber forensics process and the influences when automation is being applied to it. This view or assumptions on this topic guided the research as part of the strategy to conduct this study (Johnson & Clark, 2006). Considering the phenomena of the literature and research regarding the philosophical stances, it was found a positivistic approach was most suitable for this research. The principal beliefs of a paradigm can be summarised in three fundamental points which include ontology, epistemology, and methodology (Saunders, Lewis, & Thornhill, 2012). In the next paragraph the researcher provides an overview of the positivistic approaches as it pertains to this research. According to Saunders, Lewis and Thornhill (2019), positivism philosophical stance as follows: i) its ontological view is that there is one true reality, ii) its epistemological view is that knowledge is created from observable and measurable facts which can be used to generate law-like generalisations and iii) its axiological view is that it is a value-free research in which the researcher is detached, neutral and independent from the phenomenon being investigated.

3.3 Research design

There are three main research designs and these are the qualitative, quantitative and mixed method (Creswell and Creswell, 2018). This study took a quantitative research design which is aligned to the positivism research philosophy. Moreover, the quantitative research design "primarily follows the confirmatory scientific method because its focus is on hypothesis testing

and theory testing". In view of this, eleven hypotheses were generated from the literature review which shall be tested in this study.

3.4 Research strategy

Saunders *et al.* (2019) asserts that the selection of a research strategy is guided by the research objectives and amount of resources availability. In view of this, there are a number of research strategies may be incorporated in a research study but it should be noted that they are not mutually exclusive. However, there are six main research strategies that are used in many research studies and these are: experiment, case study, survey method, grounded theory, ethnography and archival research. These are briefly discussed in the following subsections.

3.4.1 Experiment

Typically, experiments are used to test causal relationships between a selected group of variables in a controlled environment. However, this method has been criticised owing to lack of external validity and also its inability to be applied in business and management research questions (Emmanuel, 2019).

3.4.2 Case study

The case study is an empirical investigation of a phenomenon in a real-life context which incorporates the use of multiple sources of evidence. This research study highlights the importance of context when analysing a phenomenon from a single big case study or a number of smaller cases. Furthermore, it is important to note that the boundaries between the phenomenon under investigation and the context are not usually obvious and this should be put into consideration when drawing conclusions from the findings of the study (Emmanuel, 2019).

3.4.3 Grounded theory

Saunders *et al.* (2019) contends that grounded theory is useful for research to predict and explain behaviour with the main emphasis being to develop and build theory mainly through deduction and induction. Theory is developed from data generated from a series of observations.

3.4.4 Ethnography

Ethnography is rooted firmly in the inductive approach and its purpose is to describe and explain the social world using the research subjects inhabit in the way in which they would

describe and explain it (Achari, 2014). This research strategy is time consuming and will not be used by the researcher.

3.4.5 Survey method

Emmanuel (2019) states that the survey method is a quantitative analysis research strategy that typically uses questionnaires to collect data from the respondents. However, this method is often criticised for its lack of validity due to the fact that respondents are unable to add more information as they are limited to filling in questionnaires (Fisher and Buglear, 2010). Furthermore, once the data has been collected it cannot be changed or adjusted if it is discovered that some important information was not included during the data collection process (Saunders *et al.*, 2019). Nonetheless, the survey method is still a cheap and effective way of collecting data from a large population within a reasonable timeframe. Furthermore, in the context of the Covid-19 pandemic it is a safe method as it can be conducted remotely using online surveys. Consequently, this research strategy was used in this study.

This study employed the survey method. Considering that the researcher had limited time and resources, the use of the survey method was appropriate because this method is: i) relatively inexpensive, ii) can easily be explored over a large population, iii) is flexible as it can be described using various models such as through online, email, social media, face to face and hard copy/paper surveys which can be used to collect data from remote and hard to reach participants and iv) it is also dependable as the anonymity it presents allows the participants to answer honestly.

3.5 Population of the study

A study population in research studies can be defined as the holistic assessment of the objects considering the statistical attributes estimated through sample selected for the research (Krieger, 2012). In this regard, the population of the study were cyber forensics and IT security experts within the republic of South Africa. These were chosen because they were the industry experts and they worked with cybersecurity issues on a daily basis. Consequently, they were better suited to be participants in this study due to the valuable information and experience they brought.

3.6 Sampling technique

Showkat and Parveen (2017) asserts that there are two main techniques for sampling and these are probability and non-probability sampling. Probability sampling is more associated with the positive research philosophy and the quantitative research design and because of

this it was chosen for this study. Moreover, there are five types of probability sampling techniques which are simple random sampling, stratified random sampling, cluster sampling and systematic random sampling (Showkat and Parveen, 2017). Simple random sampling was used in this study and it involves randomly selecting participants from a target population with each participant having an equal chance of being selected. In light of this, cyber forensics and IT security experts in South Africa were sampled from LinkedIn using Qualtrics.

3.6.1 Sample sizing

According to Israel (2008:1), the sample sizing is influenced by several factors which include "purpose of the study, population size, the risk of selecting a 'bad sample' and the allowable sampling error". However, the sample sizing in this study was done a simplified sample sizing formula devised by Yamane (1967):

$$n = \frac{N}{1 + Ne^2}$$

Where: N – target population size, n – sample size, e – is the level of precision which is typically 5% or 0.05.

Given that the target population of the study was 163 cyber forensics experts, the minimum sample sizing required for this study was computed below:

$$n = \frac{N}{1 + Ne^2} = \frac{163}{1 + 163(0.05^2)} = 115.8 \approx 116$$

Nevertheless, 163 cyber forensics experts were sampled and out of these 151 responded to the invitations to participate in this study. Moreover, out of the 151 participants, 122 of them managed to complete the questionnaires within the timeframe allotted to them and had indicated that they were conducting automation of cyber forensics investigation within their organisations, 13 of them indicated that they were not automating cyber forensics investigation within their organisation and 16 of them could not complete the questionnaire within the given time frame. In view of this, 13 participants who were not automating cyber forensics investigation in their organisation were not included in the sample as well those participants who did not complete their questionnaire within the allotted time. As a result, 29 questionnaires were discarded from the study and 122 participants became the sample size of the study. Consequently, the response rate of this study was 74.8% and it met the minimum required sample size of 116 as computed above.

3.7 Research instrument

According to Yaya (2014), interviews, questionnaires, observation, focus group discussion and experiment are the most common research instruments used globally. Nonetheless, this study used questionnaires as the research instrument. The research instrument was divided into four sections as illustrated in the following Table 1 below:

Table 1: Questionnaire structure breakdown

QUESTION	DESCRIPTION	REFERENCE
SECTION A	**DEMOGRAPHICS**	
A1	What is your current age group?	University of Cape Town's questionnaire policy
A2	What is your highest educational level completed?	University of Cape Town's questionnaire policy
A3	What is your current employment status?	University of Cape Town's questionnaire policy
A4	In which province(s) are you currently employed?	University of Cape Town's questionnaire policy
A5	Which best describes your current title?	University of Cape Town's questionnaire policy
A6	Which industry do you currently work in?	University of Cape Town's questionnaire policy
SECTION B	**AUTOMATION OF CYBER FORENSICS**	
B1	Do you use automation during your cyber forensics investigation process?	
B2	Do you conduct automation of cyber forensics investigation in the following phases? SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4], SA = Strongly Agree [5]	
	a. Preparation & Identification Phase	
	b. Collecting & Preservation Phase	
	c. Examination & Analysis Phase	
	d. Presentation Phase	

B3	When applying automation to the forensics investigation process, is the process faster/more efficient?	
SECTION C	**FACTORS INFLUENCING AUTOMATION OF CYBER FORENSICS INVESTIGATION**	
C1	Indicate whether you SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4], SA = Strongly Agree [5] with the following statements based on your experience within the cyber forensics investigation.	
	a. My organisation has measures to enhance accessibility of data which helps to improve automation of cyber forensics investigation during the preparation and identification phase. [Accessibility to data]	
	b. My organisation has tools and systems in place to deal with challenges associated with unstable data in order to help the automation of cyber forensics investigation during the preparation and identification phase. [Unstable data]	
	c. My organisation uses Trust Cloud to facilitate the automation of cyber forensics investigation during the collection and preservation phase. [Dependency on Cloud Service Providers (CSPs)]	
	d. My organisation uses Digital Forensics as a Service (DFaaS) to facilitate the automation of cyber forensics investigation during the collection and preservation phase. [Minimize time, maximize coverage]	

	e. My organisation uses Push Button Forensics (PBF) such as EnCase, Forensic Tool Kit, BelkaSoft and Autospy Forensic Browser to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Decline in expert knowledge]	
	f. My organisation has sufficient hardware requirements such as memory size, central processing power and disk space to enable the automation of cyber forensics investigation during the examination and analysis phase. [Hardware requirements]	
	g. My organisation has the competencies to conduct profiling and event construction to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Profiling and event construction]	
	h. My organisation has put Artificial Intelligence (AI) systems to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Artificial Intelligence]	
	i. My organisation has expert investigators that help to facilitate the automation of cyber forensics investigation during the presentation phase. [Non-expert investigator]	
	j. My organisation has tools and systems that help to collect data that has been spread over the cloud to facilitate the automation of cyber forensics	

	investigation during the presentation phase. [Spread of data in the cloud]	
	k. My organisation has put measures to overcome the reliability and privacy concerns that impedes the automation of cyber forensics investigation during the presentation phase. [Reliability and privacy]	
2	Which of the following factors to you deem the most important to your organisation?	
SECTION D	**AUTOMATION OF CYBER FORENSICS INVESTIGATION PERFORMANCE**	
D1	Indicate by whether you SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4] or SA = Strongly Agree [5] with the following statements.	
	a. Since automating cyber forensics investigation there has been a decline of cyber security incidents.	
	b. There has been an increase in the number of successful prosecutions as a result of automation of cyber forensics investigation.	
	c. Since automating cyber forensics investigation there has been an increase of productivity time and a decline in IT operational costs.	

3.8 Data collection

The study adopted a probability simple random sampling method. An online survey was created using Qualtrics, whereby it was distributed to a list of 203 contacts from an online forum for various cyber-forensic experts (including cyber-forensics investigators, cyber forensics lawyers and IT security specialists) via email. The data collection period was kept open from 22 February 2022 to 10 June 2022, during which 122 valid responses were collected.

3.9 Data analysis

The deductive approach is suitable when research is guided by existing data theory using a model with established concepts (Cho and Lee, 2014). Data analysis and collection was guided with the proposed general digital forensics model. IBM SPSS version 26 statistical package software was used to conduct the data analysis in this study. Both descriptive and inferential statistics were used in this study which included hypothesis testing, conducting correlations and regression analysis which are briefly discussed below.

3.9.1 Hypothesis testing

Generally hypothesis testing involves taking two positions the null hypothesis (H_0) and the alternative hypothesis (H_1) (Majaski, Khartit and Kvilhaug, 2021). The null hypothesis assumes that there is no difference between parameters being tested whilst the alternative hypothesis assumes there is a difference between the parameters being tested (Majaski *et al.*, 2021). In view of this, hypothesis testing allows the researchers to do inferences about the population being investigated which are then tested statistically. It is reported that hypothesis testing helps to specify the research problem being investigated, can help the researcher to know which data should be collected, improves the objectivity of the study and also helps to come up with meaningful conclusions of what can be deemed true or false. However, it should be noted that when using hypothesis testing in a study there is a possibility that mistakes can be made when drawing conclusions on these hypotheses. In light of this, there are two types of errors that are commonly experienced during hypothesis testing and these are Type I errors and Type II errors (Banerjee, Chitnis, Jadhav, Bhawalkar and Chaudhury, 2009). Type I errors are made when the study accepts the alternative hypothesis when in fact the null hypothesis is true whilst Type II error occurs when the study rejects the alternative hypothesis when in fact it is true (Banerjee *et al.*, 2009).

3.9.2 Pearson's correlation

Gupta (2016) states that correlation is a good way to indicating the relationship between two variables. However, it should be noted that correlation does not mean causation. Pearson's correlation is one of the most popular ways to compute correlations and it was computed using the following equation below (Glen, 2022):

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[\, n\sum x^2 - (\sum x)^2 \,][\, n\sum y^2 - (\sum y)^2 \,]}}$$

According to Gupta (2016), it is important to note that correlations, i) do not measure the strength of non-linear associations, ii) do not account for accidental relationships, iii) can be affected by sample bias and data contamination and iv) relationships between variables should be contextualised to avoid giving wrong conclusions. Moreover, the strength of relationships from the correlations was interpreted using the following Table 2 below:

Table 2: Correlation interpretation

Absolute Magnitude of the Observed Correlation Coefficient	Interpretation
0.00–0.10	Negligible correlation
0.10–0.39	Weak correlation
0.40–0.69	Moderate correlation
0.70–0.89	Strong correlation
0.90–1.00	Very strong correlation

Source: Schober, Boer and Schwarte (2018:1765)

3.9.3 Regression analysis

Regression analysis is a statistical modelling technique that tries to creates a linear relationship between the dependent variable and independent variables by fitting the observed data points on a linear equation which the significant and appropriateness of fit can check statistical outputs such as scatter plots (Gupta, 2016). According to Moore, Anderson, Das and Wong (2006:235), “Multiple regression is a statistical technique that can be used to analyse the relationship between a single dependent variable and several independent variables. The objective of multiple regression analysis is to use the independent variables whose values are known to predict the value of the single dependent value. Each predictor value is weighed, the weights denoting their relative contribution to the overall prediction”. In this regard, examining the predictor value for each factor (see Section 2.8) that influences the automation of cyber security investigation performance may help to understand the influence of each independent variable (i.e., Accessibility to data, Unstable data, Dependency CSPs, DFaaS, PBF, Hardware requirements, Profiling and event reconstruction, AI, Non-expert investigator, Spread of data and Reliability and privacy) on the dependent variable (i.e., Automation of cyber forensics investigation performance). However, the linear regression equation is typically represented by the following equation:

$$Y = a + bX$$

Where: X – independent variable, Y – dependent variable, b – gradient and a – y-intercept.

3.10 Validity and reliability

Validity of a research study is defined as "the extent to which a concept is accurately measured in a quantitative study" (Heale and Twycross, 2015:66). The study ensures validity of the research findings by conducting a pilot study. This allowed the study to identify any potential issues with the research instrument before the study was conducted. In view of this, five cyber forensics experts were asked to participate in the pilot study. However, no major issues apart from grammatical issues were identified from the research instrument and were rectified accordingly.

On the other hand, reliability of a research study "relates to the consistency of a measure" (Heale and Twycross, 2015:66). There are four main threats to research reliability and these are participant bias, participant error, observer bias and observer error (Robson, 2002). In view of this, internal consistency of research instruments was checked using Cronbach's alpha test to ensure reliability of the findings. Furthermore, the research findings were triangulated with existing literature and were also moderated and contextualised to the research study. Additionally, the researcher did not influence the responses of the participants in any way and the data collection and analysis were done objectivity. Furthermore, participants error and bias were moderated by using a large sample size and appropriate statistical tools.

3.11 Ethical considerations

Ethical approval for the research was obtained from the University of Cape Town Ethics Committee. The research invitation email contained all necessary information regarding the aim and limitations of the research, and confidentiality was assured by removing personally identifiable information from the dataset. The researcher secured informed consent from all participants in this research.

Completion of the survey was voluntary, with no repercussions. All responses were kept securely in a password protected account that is only accessible by the researcher. The possibility of making some results of digital investigations public, might infringe the ethical policy. No identifiable information is mentioned, in order to maintain confidentiality.

The structure of the research study gives context to how the investigators go about their process and incorporating automation tools to achieve the best result. While no predetermined

hypotheses are pursued, results provide a more structured view of factors that should be considered in the practice of cyber forensic science.

3.12 Chapter summary

The study took a positivism research philosophy using the quantitative research design. Furthermore, probability simple random sampling was used for this study and the sample size for this study was 122 participants. The data was analysed using IBM SPSS version 26 to compute both descriptive and inferential statistics. The next chapter discusses the research results of the study.

Chapter 4: DATA ANALYSIS & DISCUSSION

4.1 Introduction

This chapter outlines the data analysis and discussion of the study. It shall be divided into five main sections which are demographics of the participants, automation of cyber forensics, factors influencing automation of cyber forensics, cyber forensics investigation performance and correlations. These shall be discussed in detail in the following sections.

4.2 Demographics of the participants

This section discusses the findings from the demographics of the participants in this study. It will be divided into seven subsections which are age, education, employment status, province, job title, tenure and industry.

4.2.1 Age

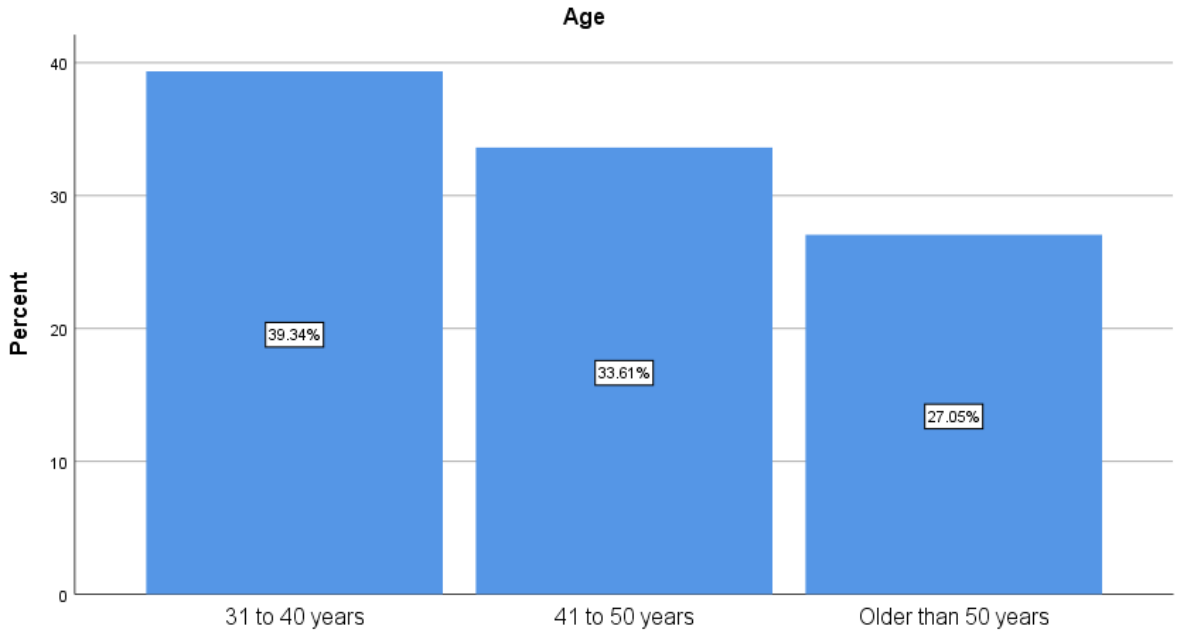

Figure 3: Age of participants

Most of the participants were aged 31 to 40 years old and these accounted for 39.34% of the participants as shown in Figure 3 above. However, the rest of the participants were aged 41 to 50 years and older and these accounted for 33.61% and 27.05% of the participants respectively. These findings suggests that the participants were mature which could positively influence the quality of their responses as older people tend to be more logical and analytical.

4.2.2 Education

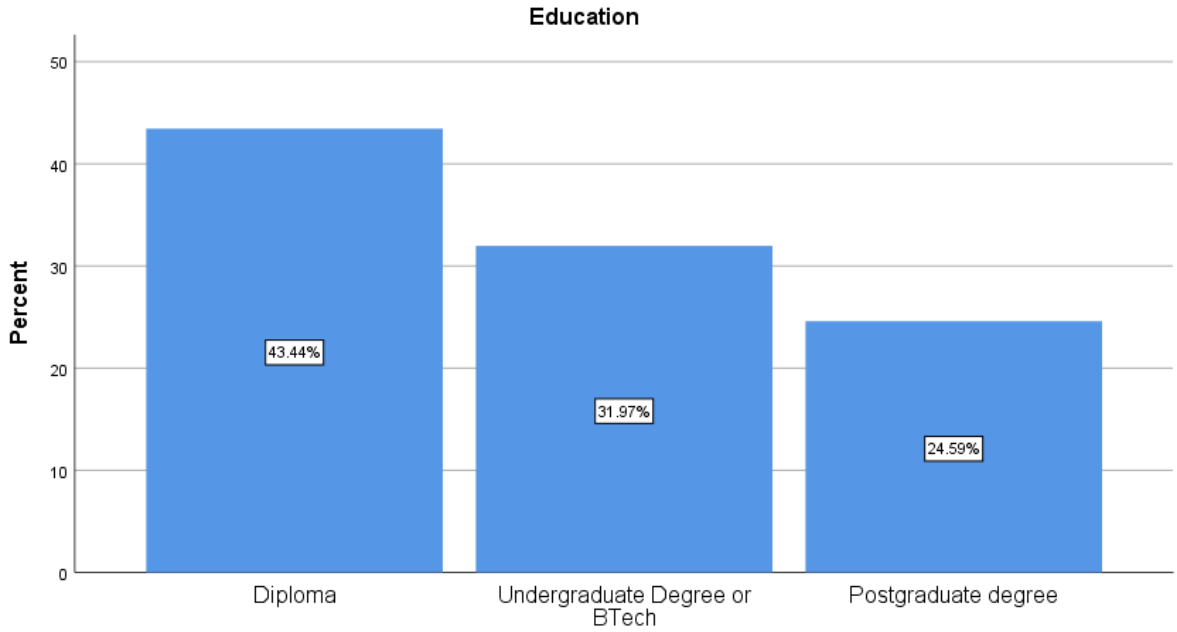

Figure 4: Education level of participants

Most of the participants in this study had a diploma as their highest level of education and this accounted for 43.44% of the participants as shown in Figure 4 above. Nevertheless, the rest of the participants had an undergraduate degree or BTech and postgraduate degree and these accounted for 31.97% and 24.59% of the participants respectively. In view of this, it can be inferred that the level of education of the participants was relatively high and that they were in a good position to contribute positively to this study.

4.2.3 Employment status

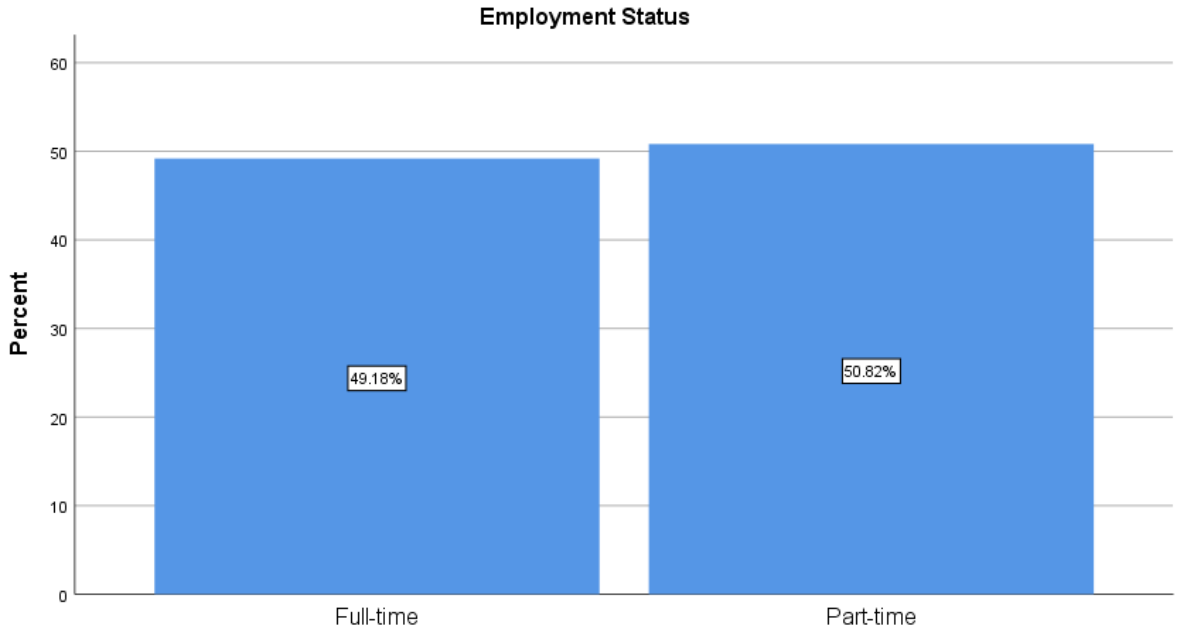

Figure 5: Employment status of participants

The majority of the participants indicated that they were working part-time and these accounted for 50.82% of the participants. Nonetheless, the rest of the participants indicated that they were working full-time and this accounted for 49.18% of the participants as shown in Figure 5 above. The sheer number of participants working part-time could have been influenced by the recent Covid-19 pandemic which made many organisations cut their labour costs and where they made some roles part time to save on labour costs.

4.2.4 Province

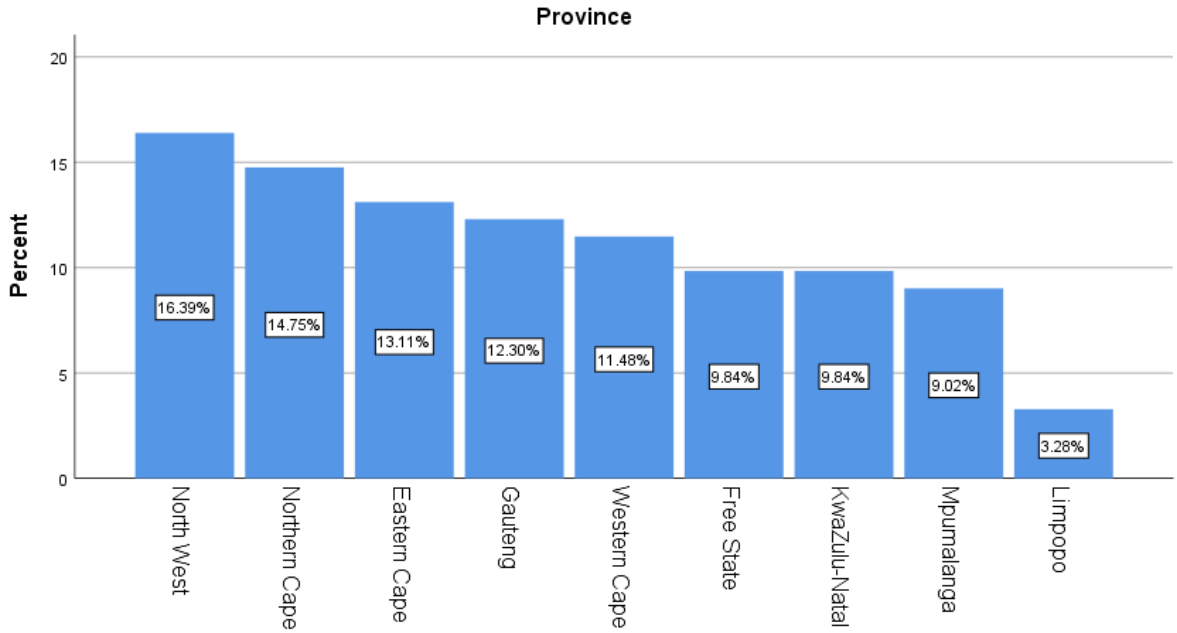

Figure 6: Province in which participants are located

The findings in Figure 6 indicate that most of the participants were from North-West province and these were closely followed by those from the Northern Cape. These two accounted for the 16.39% and 14.75% of the participants respectively as shown in Figure 5 above. However, the rest of the participants were from Eastern Cape, Gauteng, Western Cape, Free State, KwaZulu-Natal, Mpumalanga and Limpopo provinces and these accounted for 13.11%, 12.30%, 11.48%, 9.84%, 9.84%, 9.02% and 3.28% respectively. These findings suggests that all nine provinces were fairly represented in this study and this could help to gain an understanding of how cyber forensics investigation are being conducted in different regions of the country.

4.2.5 Job title

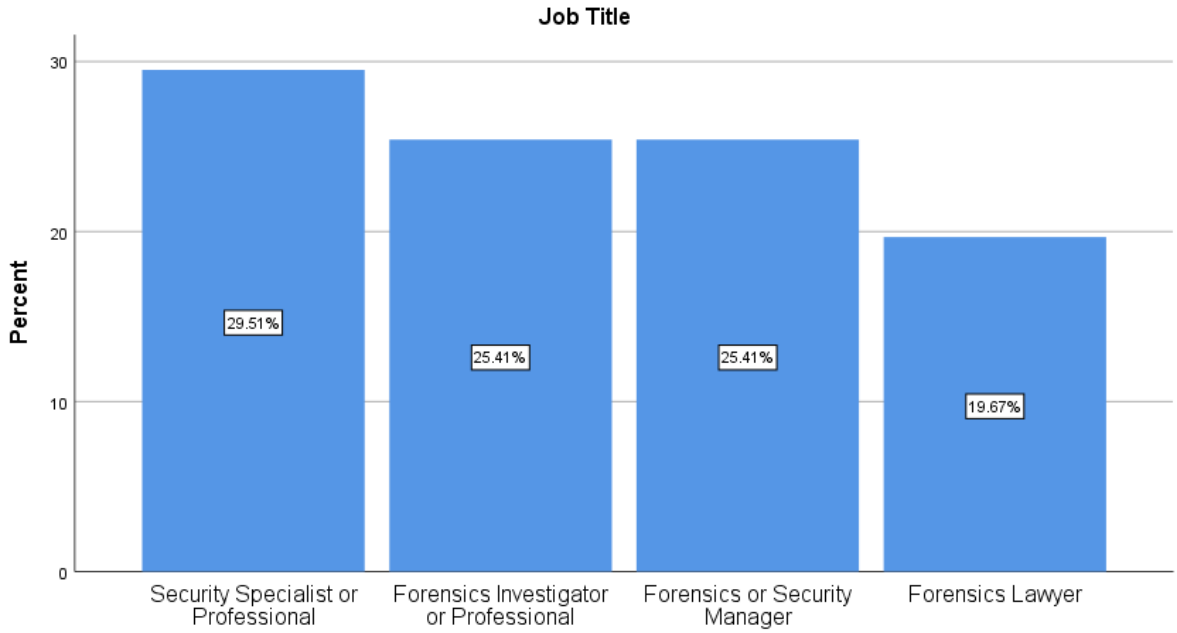

Figure 7: Job title of participants

Most of the participants in this study were security specialists or professionals and these accounted for 29.51% of participants as shown in Figure 6 above. These were closely followed by those who were forensics investigators or professionals and forensics or security managers and these accounted for 25.41% each of the participants in this study. Nevertheless, 19.67% of the participants were forensics lawyers as shown in Figure 7 above. These findings suggests that the participants were fairly balanced across difference disciplines of the cyber forensics investigation areas. This could help to bring out important and rich findings in this study due to different experiences, perspectives and opinions of the experts that participated.

4.2.6 Tenure

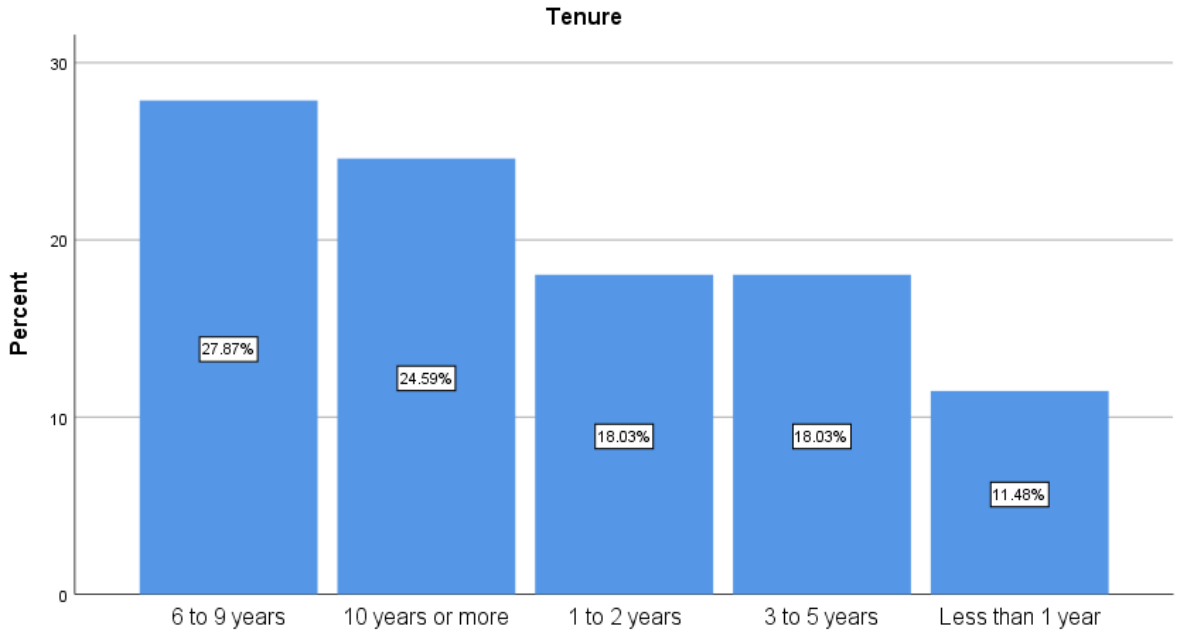

Figure 8: Tenure of participants in their current organisation

Most of the participants had 6 to 9 years tenure within their current organisation and this accounted for 27.87% of the participants. This was closely followed by those who had at least 10 years tenure within their current organisation and these accounted for 24.59% of the participants. Nevertheless, those who were ranked third in their tenure had 1 to 2 years and 3 to 5 years and these each accounted for 18.03% of the participants respectively as shown in Figure 8 above. It had been noted from Figure 8 that the majority of the participants had at least 6 years tenure within their current organisations and this accounted for 52.46% of the participants. This suggests that the majority participants had enough work experience with the cyber forensics investigation within their current organisation which could have yielded quality responses from the participants.

4.2.7 Industry

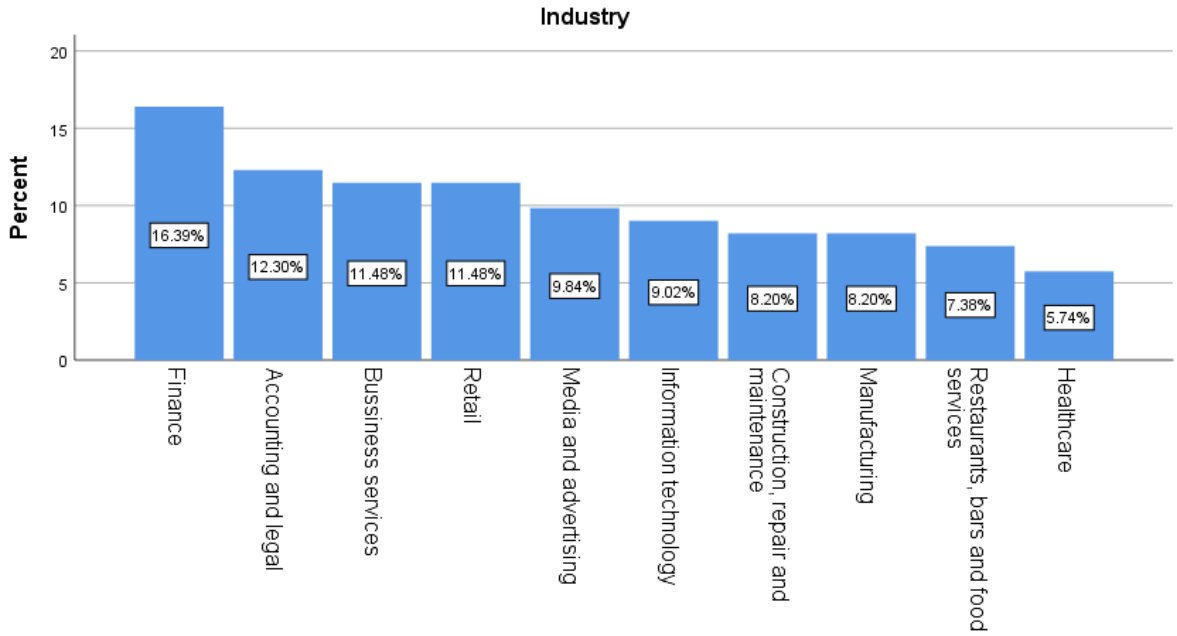

Figure 9: Industry that participants work in

Most of the participants in this study were in the finance industry as shown in Figure 9 above. However, other industries were also fairly represented in this study and these include the accounting and legal, business services, retail, media and advertising, information technology, construction, repair and maintenance, manufacturing, restaurants, bars and food services and healthcare. These accounted for 12.30%, 11.48%, 11.48%, 9.84%, 9.02%, 8.20%, 7.38% and 5.74% respectively. These findings suggest that the participants were from a wide array of industries which could have helped to bring different experiences of the participants on cyber forensics investigation in relation to what they encounter in their respective industry.

4.3 Automation of cyber forensics

This section discusses automation of cyber forensics investigation. It is divided into six subsections which are automation of cyber forensics investigation process, automation of preparation and identification phase, automation of collection and preservation phase, automation of examination and analysis phase, automation of presentation phase and improvement of cyber forensics investigation. These are shown in Table 3 and discussed in detail in the following subsections.

Table 3: Questions of automation of cyber forensics

Do you use automation during your cyber forensics investigation process?		**Yes**		**No**	
		100%		0%	
Do you conduct automation of cyber forensics investigation in the follows phases? **SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4], SA = Strongly Agree [5]**	**SD**	**D**	**N**	**A**	**SA**
a. Preparation & Identification Phase	9.84%	9.84%	7.38%	31.97%	40.98%
b. Collecting & Preservation Phase	0.82%	0%	0%	45.90%	53.28%
c. Examination & Analysis Phase	28.69%	45.08%	12.30%	13.93%	0%
d. Presentation Phase	8.20%	64.75%	7.38%	13.11%	6.56%
When applying automation to the cyber forensics investigation process, is the process faster or more efficient?		**Yes**		**No**	
		50%		50%	

All the participants indicated that they conducted automation of cyber forensics investigation (100%). This suggests that the participants perform some form of automation of cyber forensics investigation process within their organisations. Moreover, by the participants indicating that they were performing automation of cyber forensics investigation process within their organisation implies that they had experience in this regard which could lead to meaningful contribution to the study. The importance of the automation of cyber forensics investigation is increasingly becoming more important due to the emergence of big data and because of this, automation can help to increase the speed, accuracy and effectiveness of the cyber forensics investigation process (Avian, 2022).

The majority of the participants indicated that they automated the preparation and identification phase of the cyber forensics investigation and this was supported by a total agreeableness score (agree + strongly agree % scores) of 72.95% as shown in Table 3 above. However, some of the participants indicated that they did not automate the preparation and identification phase within their organisation and this is supported by a total disagreeableness score (disagree + strongly disagree % scores) of 19.68%. Nevertheless, 7.38% of the participants

were neutral on this matter. These findings suggests that the majority of the organisations are automating preparation and identification phase of their cyber forensics investigations. The preparation and identification phase of the cyber forensics investigation process is primarily influenced by the accessibility of data and unstable data. According to Rane and Dixit (2019), tools such as Syslog and Log Analyzer can facilitate the automation of cyber forensics investigation process by enhancing accessibility to data. In this regard, participants who have appropriate tools within their organisation are able to automate the cyber forensics investigation and those who do not have are therefore to do so as shown in the results above. On the other hand, unstable data can hamper the automation process (MacDermott *et al.*, 2018) and may therefore require human interaction to access and/or protect the data (Damshenas *et al.*, 2012). As a result, organisations that deal with unstable data as part of their cyber forensics investigation process would be difficult relative to those who do not deal with unstable data. This may explain why some of the participants indicated that they do not automate the preparation and identification phase of the cyber forensics investigation process.

The majority of the participants indicated that they were automating their collection and preservation phase of the cyber forensics investigation within their current organisations. This was supported by a total agreeableness score of 99.18% as shown in Table 3 above. On the other hand, only 0.82% of the participants indicated that they did not automate their collection and preservation phase of the cyber forensics investigation. It seems automation of the collection and preservation phase is being done in many of the organisations which were presented in this study. It is also important to note that the collection and preservation phase of the automation cyber forensics investigation is influenced by the dependency on CSPs and DFaaS. CSPs when there is a breach from a cycle attack, they typically prioritise service restoration than data protection in their cloud system design as would want to make sure their clients are happy. But this makes it harder for the cyber forensics investigations to do their job. Additionally, Morales-Ferreira *et al.* (2018) stated that the CSPs may have several data centres distributed globally in undisclosed locations. This also hinders the automation of cyber forensics investigation. However, Alqahtany (2017) states that Trust Cloud can be used as IaaS to facilitate the automation of cyber forensics investigation.

The majority of the participants in this study indicated that they did not automate the examination and analysis phase of their cyber forensics investigation process and this was supported by a total disagreeableness score of 73.77% as shown in Table 3 above. On the other hand, 13.93% of the participants indicated that they did automate the examination and analysis phase of their cyber forensics investigation process and this was supported by a total agreeableness score of 13.93%. Nevertheless, 12.30% of the participants were neutral on this

matter. These findings suggests that many of the organisations in this study did not automate the examination and analysis phase of their cyber forensics' investigation process. The examination and analysis phase of the automation of the cyber forensics' investigation process is influenced by the PBFs, hardware requirements, profiling & event reconstruction and AI. According to Horsman (2020), PBF suites such EnCase, Forensic Tool Kit, BelkaSoft and Autopsy Forensic Browser help to facilitate the automation process at the examination and analysis phase of the cyber forensics investigation. Moreover, as per the hardware requirements, there is a need of a powerful computer capable of handling big data and meeting the automation performance requirement (Homem, 2016; Irons and Lallie, 2014). Furthermore, Al Mutawa *et al.* (2019) and Horsman (2020) points out that the automation of the cyber forensics investigation process should be validated and evaluated through the utilisation of computer analysis such as profiling and event reconstruction. Also, Wylot *et al.* (2018) highlights the importance of AI in the automation process to enhance the quality of outputs of intelligence forensics. Nevertheless, it seems from the results in Table 3 that the majority of the participants' organisations are not actively automating the examination and analysis phase of the cyber forensics investigation process. This could be due to the limited resources and a low return on investment on this phase of the cyber forensics investigation process.

The majority of participants indicated that they were not automating the presentation phase of their cyber forensics investigation process and this was supported by a total disagreeableness score of 72.95% as shown in Table 3 above. On the other hand, some of the participants indicated that they were automating the presentation phase and this was supported by a total agreeableness score of 19.67%. Nonetheless, 7.38% of the participants were neutral in this matter. The findings suggests that many organisations are not automating the presentation phase of the cyber forensics investigation process. The automation of the presentation phase of the cyber forensics investigation is influenced by existence of a non-expert investigator, spread of data in the cloud and reliability & privacy. Chhabra *et al.* (2020) points out that the presentation phrase of the cyber forensics investigation process is very important as it determines what and how evidence will be presented in the court of law which could make or break the case. For instance, evidence presented by a non-expert investigator may be dismissed in court as the evidence would lose credibility. Also, the automation of the cyber forensics investigation process is a complex things and it may be difficult to explain the technicalities involved in a manner that is understandable to the jury (Simou *et al.*, 2016). This may be further complicated if the data is spread over different jurisdictions and the applicable laws may vary on how to handle and process the data. This may hamper the automation of the data. In view of this, Hughes and Karabiyik (2020) states that the reliability of the forensic

tool being employed and its ability to protect the privacy of data. Organisation that may not be able to meet at least one of those requirements may be able to automate the presentation phase of the cyber forensics investigation process. This may be the reason why the majority of the participants indicated that their organisations are not or are unable to automate the presentation phase.

The results in Table 3 above, show that 50.00% of the participants experienced faster or more efficient cyber forensics investigation process after automating it and other 50.00% did not experience any changes. This suggests that automation of the cyber forensics investigation process does not always yield positive results, and this may also imply that there are other factors that must be considered to optimise the impact of the automation process. It should be noted that the automation of the cyber forensics investigation is not only a technological issue but it also encompasses political and social issues (James and Gladyshev, 2013). In light of this, automation of the cyber forensics investigation process does not always lead to positive or expected outcomes due to a combination of technological, political and social issues that may impede the automation process.

4.4 Factors affecting automation of cyber forensics investigation.

This section discusses the factors affecting automation of cyber forensics investigation. It shall be divided into twelve subsections which are 1) accessibility to data, 2) unstable data, 3) dependency on CSPs, 4) minimise time, maximise coverage, 5) decline in expert knowledge, 6) hardware requirements, 7) profiling and event construction, 8) artificial intelligence, 9) non-expert investigator, 10) spread of data in the cloud, 11) reliability and privacy, 12) most important factor affecting the automation of cyber forensics investigation.

4.4.1 Accessibility to data

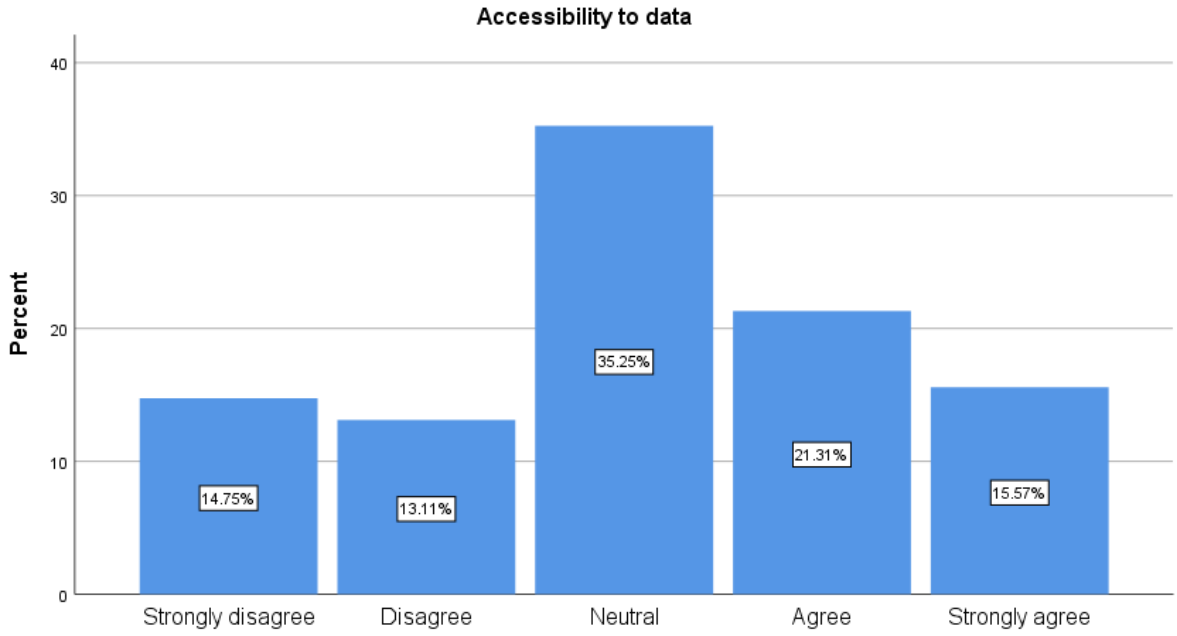

Figure 10: Accessibility to data

Most of the participants indicated that they had put measures in place to enhance accessibility to data in their organisation in order to improve the automation of cyber forensics investigation during the preparation and identification phase. This was supported by the total agreeableness score of 36.88%. On the other hand, some of the participants indicated that they had not put measures in place to enhance accessibility to data in their organisation and this was supported by a total disagreeableness score of 27.86% as shown in Figure 10 above. However, 35.25% of the participants had a neutral stance on this matter. These findings suggests that many organisations are not putting in enough effort and attention into increasing or improving the accessibility to data. This may have negative consequences when it comes to automation of the cyber forensics' investigation process within their respective organisations.

4.4.2 Unstable data

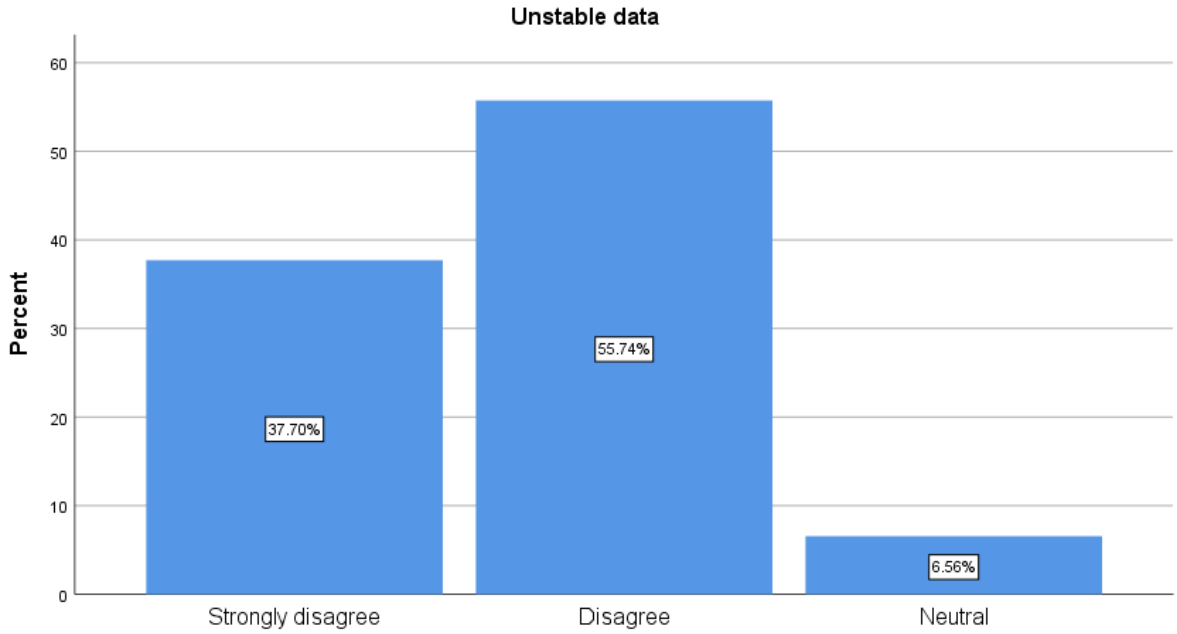

Figure 11: Unstable data

The majority of the participants indicated that their organisation did not have tools and systems in place to deal with the challenges associated with unstable data that would help to automate the cyber forensics investigation during the preparation and identification phase. This was supported by a total disagreeableness score of 93.44% as shown in Figure 11 above. Nonetheless, the rest of the participants had a neutral stance on this matter and these accounted for 6.56% of the participants. The results in Figure 11 above suggests that organisations should invest more in tools that helps to mitigate the challenges posed by unstable data. This would greatly help to improve the automation of cyber forensics investigation during the preparation and identification phase.

4.4.3 Dependency on Cloud Service Providers

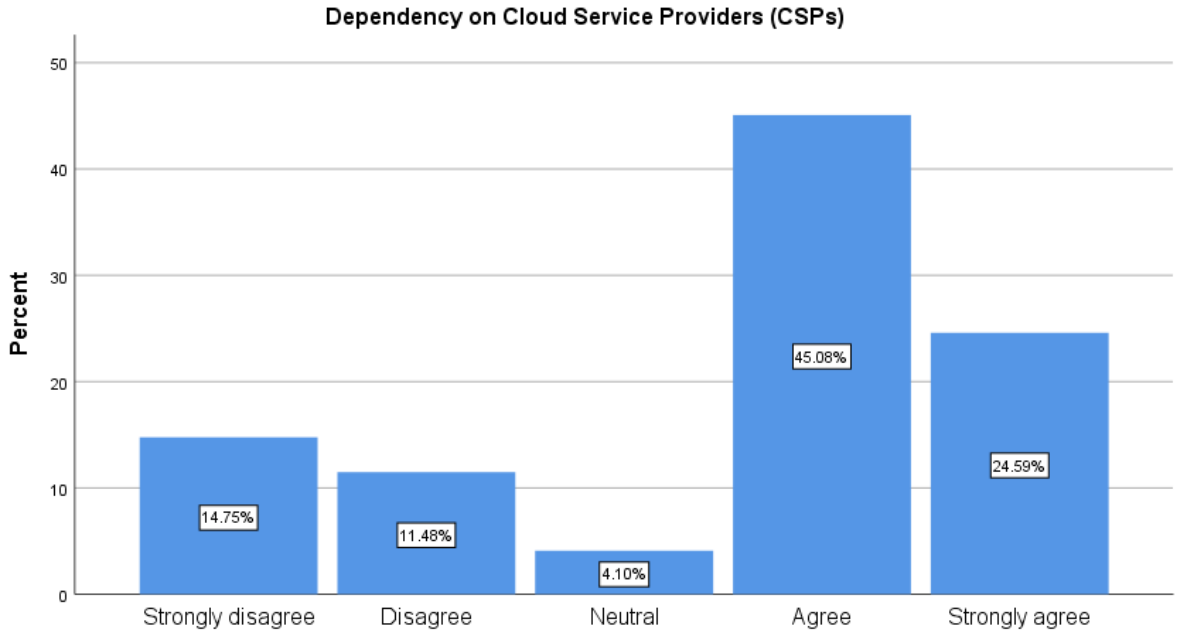

Figure 12: Dependency on CSPs

Figure 12 above indicated that the majority of the participants used Trust Cloud to facilitate the automation of cyber forensics investigation during the collection and preservation phase. This was supported by a total agreeableness score of 69.67%. On the other hand, only 26.43% accounted for the disagreeableness score of those participants who indicated that they did not use Trust Cloud within their organisation. Yet still, 4.10% of the participants had a neutral stance on this matter. These findings suggest that most of the organisation use Trust Cloud in order to overcome the challenges associated with dependency on Cloud Service Providers (CSPs).

4.4.4 Minimise time, maximise coverage

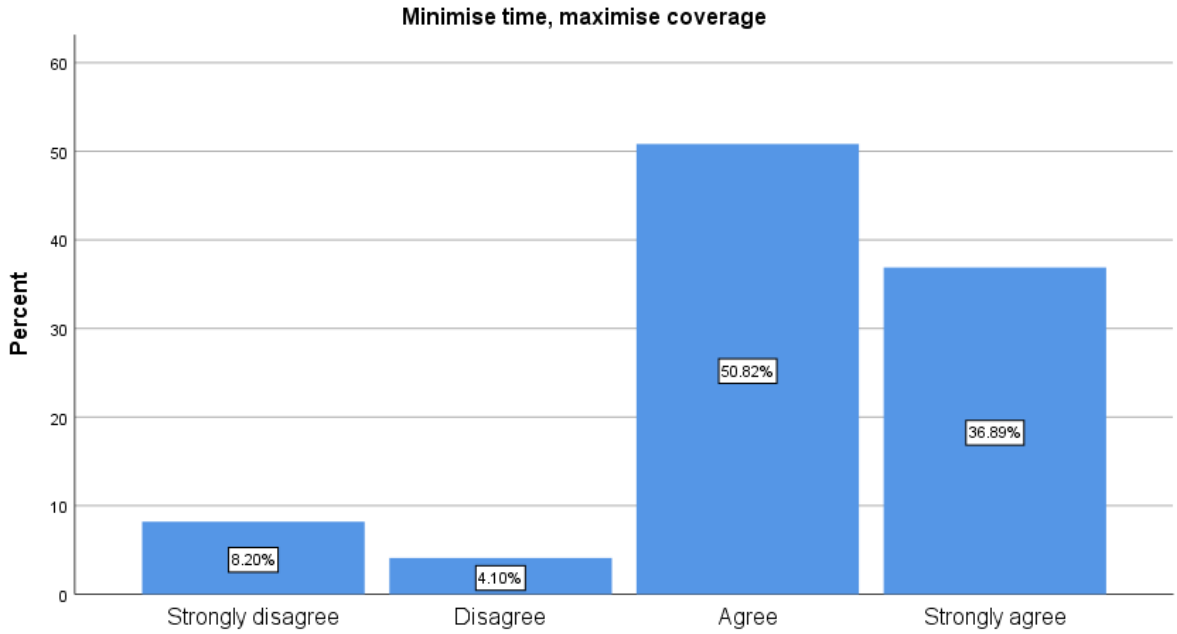

Figure 13: Minimise time, maximise coverage

The majority of the participants indicated that they used Digital Forensics as a Service (DFaaS) to facilitate the automation of cyber forensics investigation during the collection and preservation phase. This was supported by a total agreeableness score of 87.71% as shown in Figure 13 above. However, some of the participants indicated that they did not use DFaaS within their organisation and this accounted for a total disagreeableness score of 12.30%. These findings suggests that many of the organisations in this study were using DFaaS as a means to minimise time and maximise coverage which could help to facilitate the automation of cyber forensics investigation during the collection and preservation phase.

4.4.5 Decline in expert knowledge

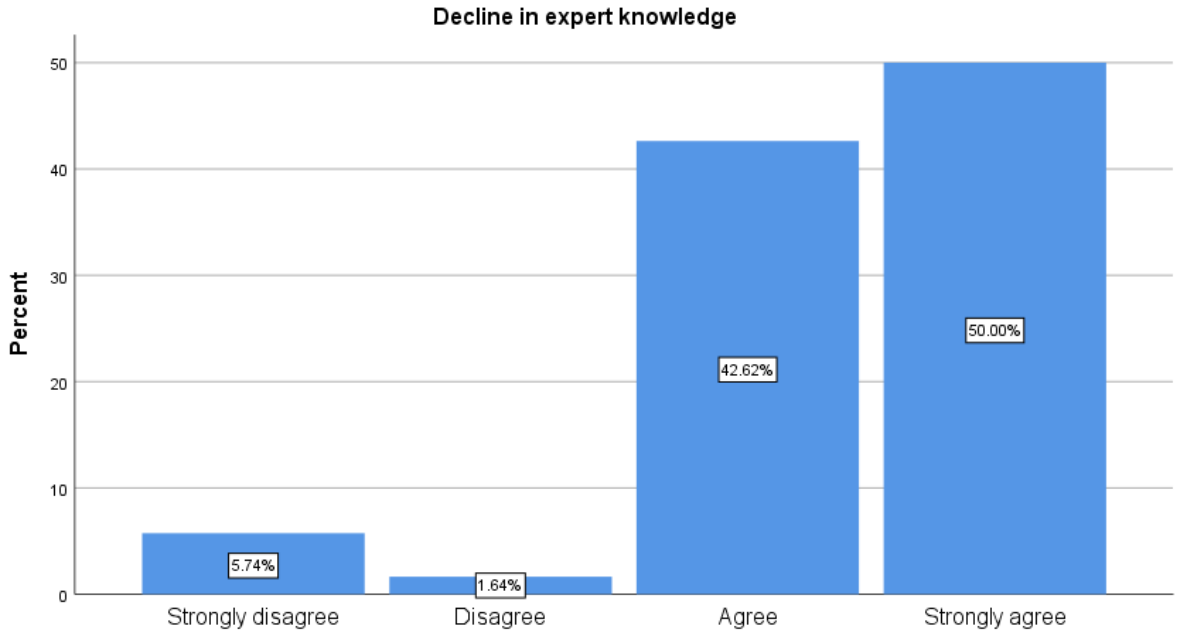

Figure 14: Decline in expert knowledge

The majority of the participants indicated that their organisations were using Push Button Forensics (PBF) such as EnCase, Forensic Tool Kit, BelkaSoft and Autospy Forensic Browser to facilitate the automation of cyber forensics investigation during the examination and analysis phase. This accounted for a total agreeableness score of 92.62% as shown in Figure 14 above. Yet still, some of the participants indicated that they did not use PBF within their organisations and this accounted for a total disagreeableness score of 7.38%. These findings suggest that the majority of the organisations in this study were using PBF to mitigate the challenges associated with the decline in expert knowledge that would make it difficult to automate the examination and analysis phase of the cyber forensics investigation process.

4.4.6 Hardware requirements

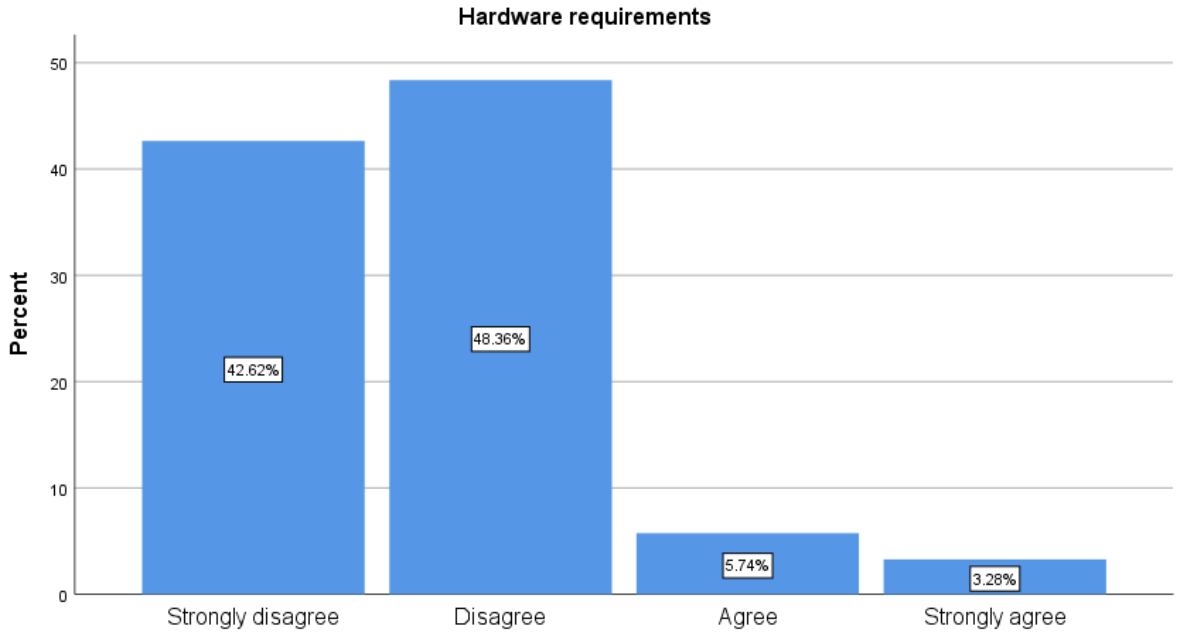

Figure 15: Hardware requirements

The majority of the participants indicated that they did not have sufficient hardware requirements such as memory size, central processing power and disk space to enable the automation of cyber forensics investigation during the examination and analysis phase. This accounted for a total disagreeableness score of 90.98% as shown in Figure 15 above. On the other hand, some of the participants indicated that they had sufficient hardware requirements and these accounted for a total agreeableness score of 9.02%. These findings suggest that the majority of the organisations in this study did not have sufficient hardware requirements which could lead to challenges when it comes to automation of the examination and analysis phase during the cyber forensics' investigation phase.

4.4.7 Profiling and event construction

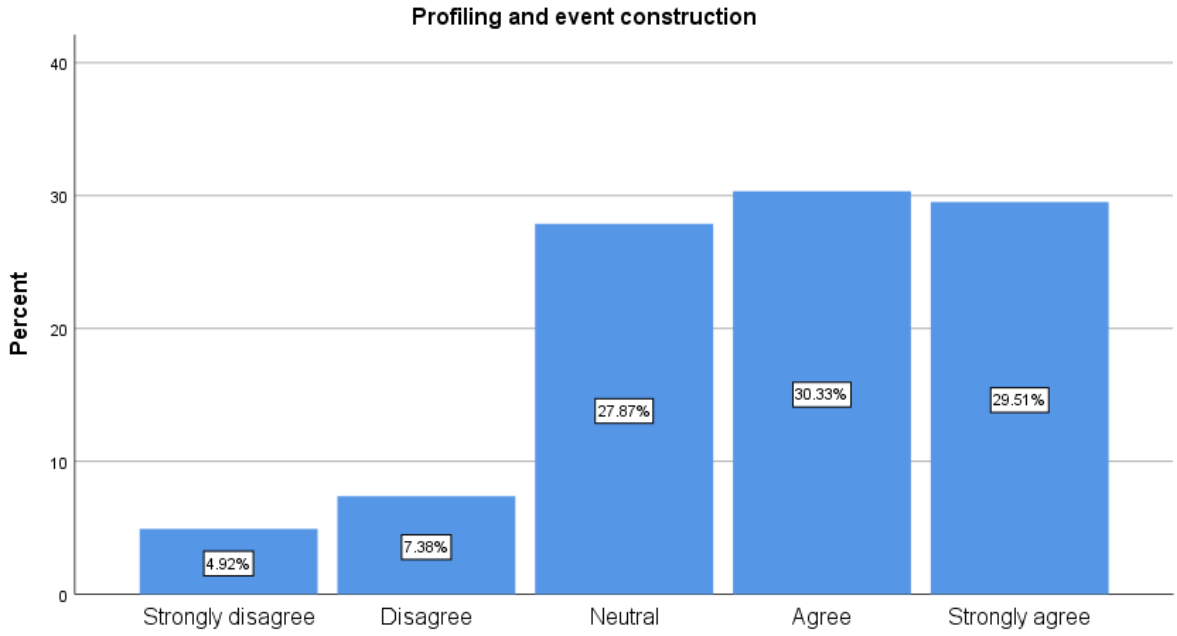

Figure 16: Profiling and event construction

The majority of the participants indicated that their organisations had the competencies to conduct profiling and event construction to facilitate the automation of cyber forensics investigation during the examination and analysis phase. This accounted for a total agreeableness score of 59.84% as shown in Figure 16 above. On the other hand, only 12.30% accounted for the total disagreeableness score as indicated by some of the organisations that did not have enough competencies to conduct profiling and event construction to facilitate the automation of cyber forensics investigation during the examination and analysis phase. Nevertheless, 27.87% of the participants had a neutral stance on this matter as shown in Figure 16 above. These findings suggest that many organisations have enough competencies to conduct profiling and event construction which is beneficial to the automation of the cyber forensics investigation process.

4.4.8 Artificial intelligence

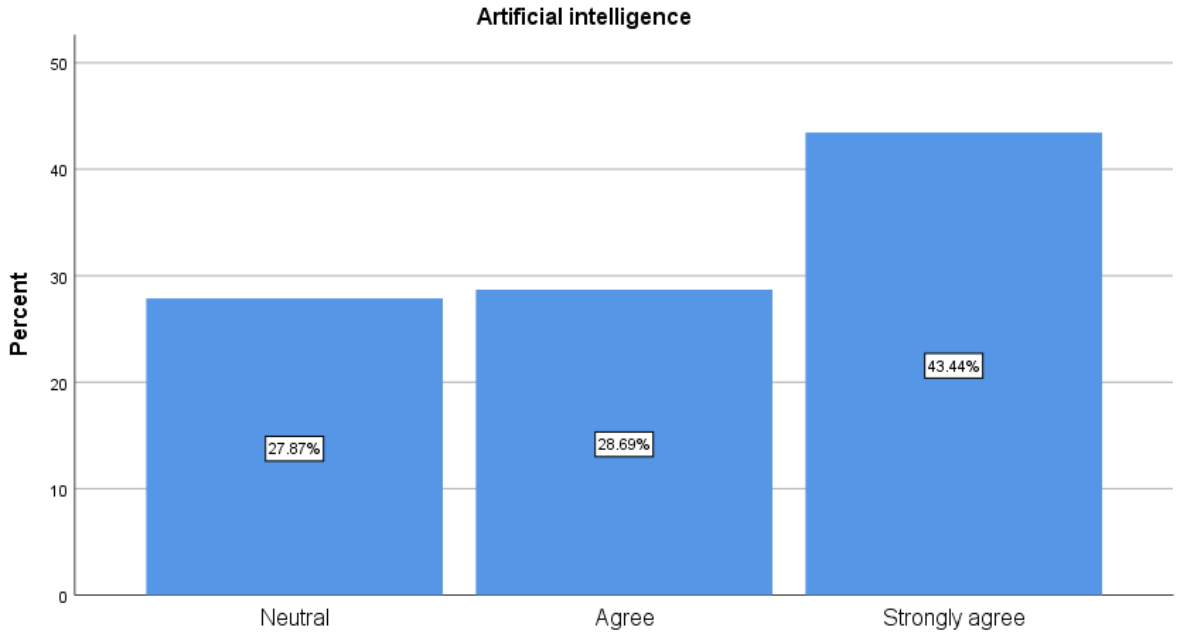

Figure 17: Artificial intelligence

The majority of the participants indicated that their organisation had incorporated artificial intelligence (AI) systems to facilitate the automation of cyber forensics investigation during the examination and analysis phase. This accounted for a total agreeableness of 72.13% as shown in Figure 17. Nonetheless, 27.87% of the participants indicated that they were neutral on this matter. These findings suggests that the majority of the organisations that were represented in this study had made considerable investments in putting AI systems in place that would help in the automation of cyber forensics investigation during the examination and analysis phase.

4.4.9 Non-expert investigator

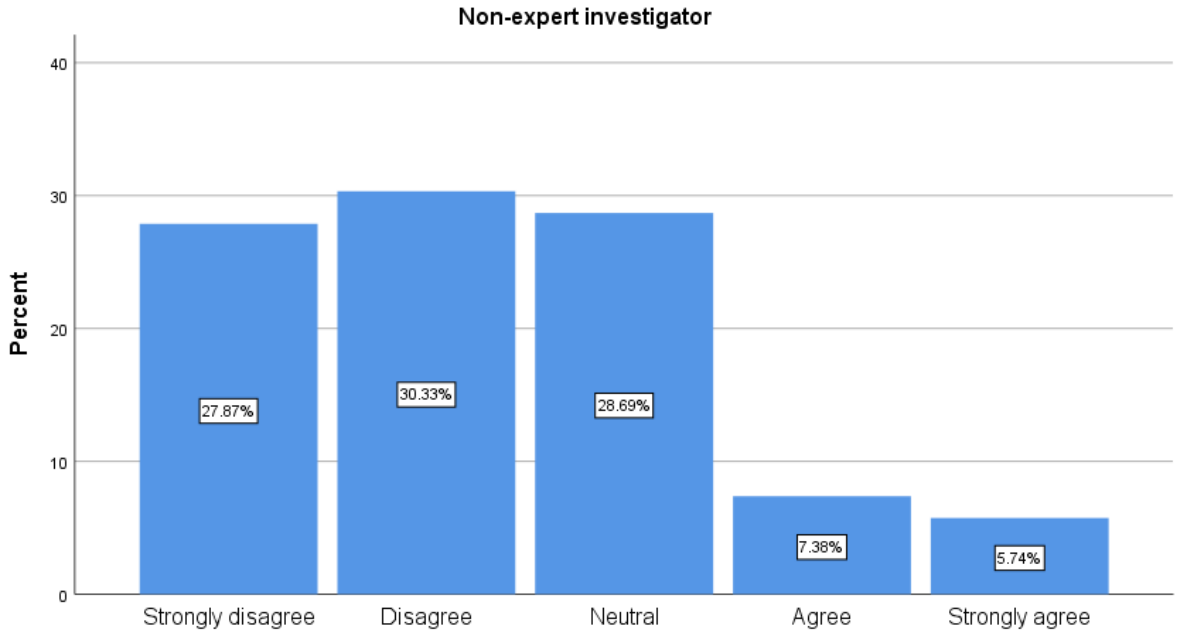

Figure 18: Non-expert investigator

The majority of the participants indicated that their organisations did not have sufficient expert investigators that help to facilitate the automation of cyber forensics investigation during the presentation phase. This accounted for a total disagreeableness score of 58.20% as shown in Figure 18 above. However, some of the participants indicated that they had sufficient expert investigators within their organisation and this accounted for a total agreeableness score of 13.12%. It should also be noted that 28.69% of the participants were neutral on this matter. Nonetheless, it can be inferred from these findings that most of the organisations that were represented in this study did not have sufficient expert investigators to help them to facilitate the automation of the cyber forensics investigation during the presentation phase.

4.4.10 Spread of data in the cloud

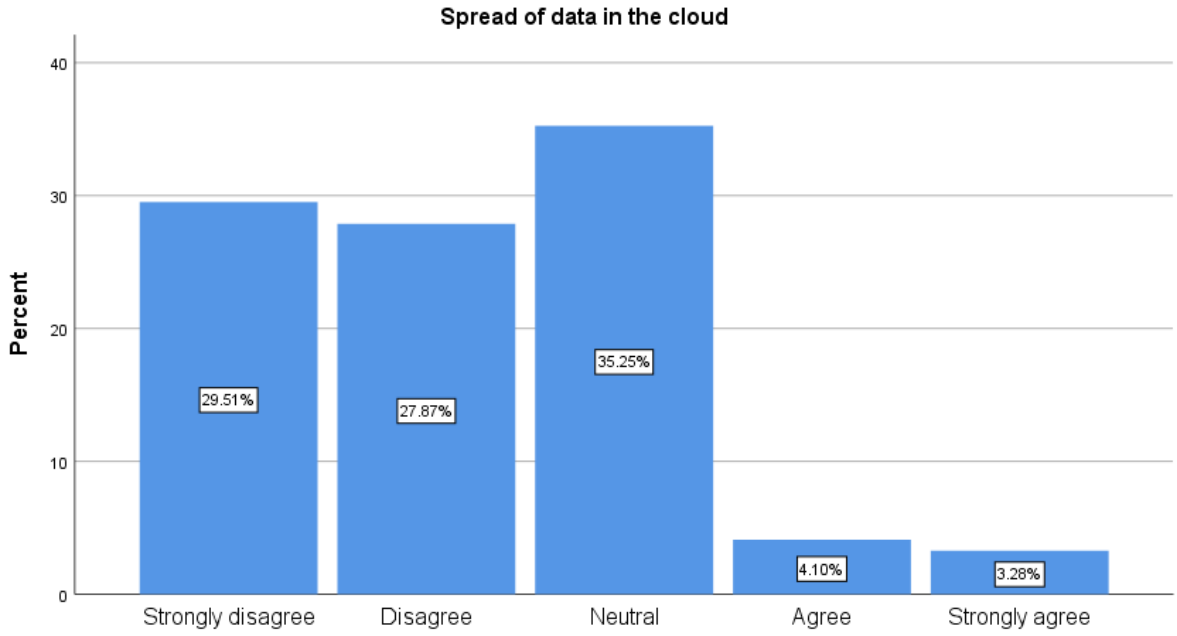

Figure 19: Spread of data in the cloud

The majority of the participants indicated that their organisations did not have tools and systems to help them to collect data that has been spread over the cloud to facilitate the automation of cyber forensics investigation during the presentation phase. This accounted for a total disagreeableness score of 57.38% as shown in Figure 19. On the other hand, some of the participants indicated that their organisations had tools and systems in place to facilitate the automation of cyber forensics investigation during the presentation phase and this was supported by a total agreeableness score of 7.38%. Yet still, 35.25% of the participants were neutral on this matter. These findings suggest that the spread of data in the cloud is a big challenge that impedes the automation of cyber forensics investigation and needs to be mitigated to improve automation during the presentation phase.

4.4.11 Reliability and privacy

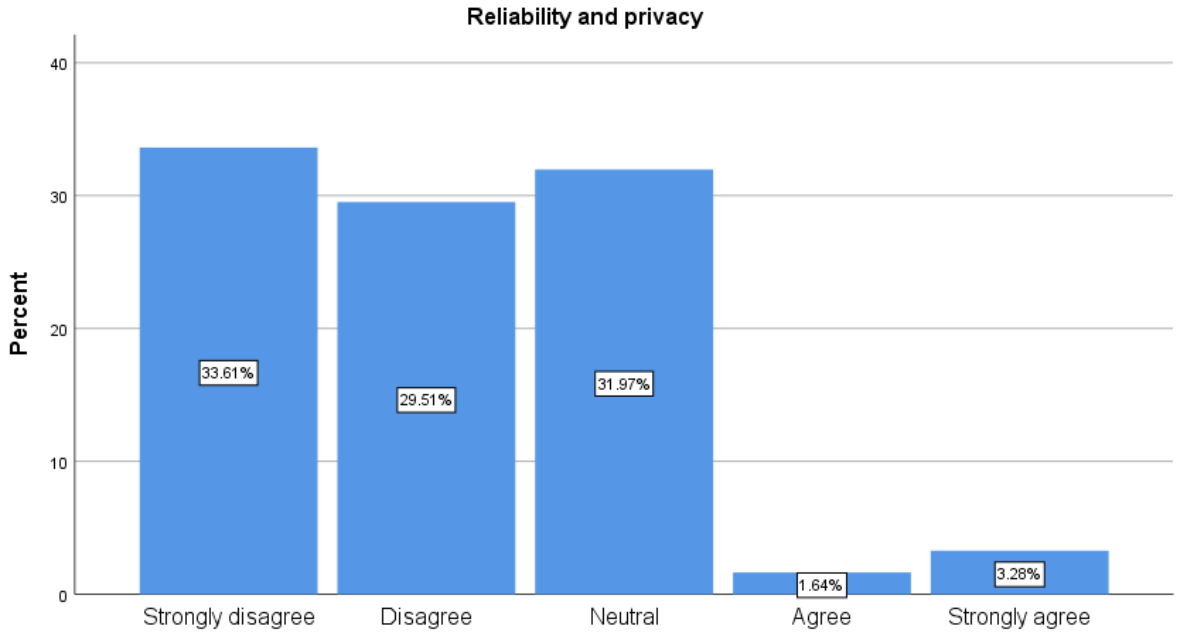

Figure 20: Reliability and privacy

The majority of the participants indicated that their organisations did not have measures put in place to overcome the reliability and privacy concerns that impedes the automation of cyber investigation during the presentation phase. This was supported by a total disagreeableness score of 63.12% as shown in Figure 20 above. Moreover, a minute portion of the participants indicated that their organisations had put measures in place to overcome the reliability and privacy concerns which accounted for a total agreeableness score of 4.92%. However, 31.97% of the participants indicated that there were neutral on this matter.

4.4.12 Most important factor

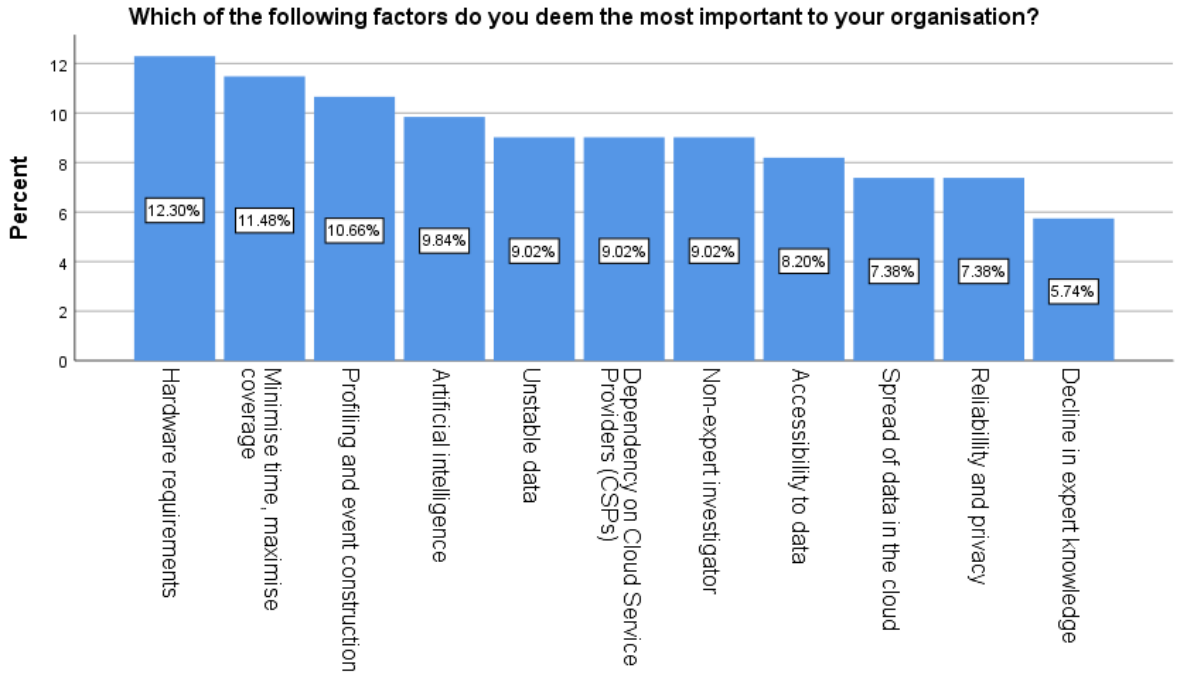

Figure 21: Most important factor for cyber forensics investigation for the participants' organisations

The results in Figure 21 above show that the top three factors that were indicated by the participants to be the most important to their organisations were 1) hardware requirements, 2) minimise time, maximise coverage and 3) profiling and event construction. These accounted for 12.30%, 11.48% and 10.66% respectively. Furthermore, some participants indicated that other factors were the ones which they deemed important to their organisations. In view of this, the fourth most important factor was artificial intelligence which accounted for 9.84% of the participants as shown in Figure 21 above. Moreover, the fifth most important factor had three factors which are unstable data, dependency on Cloud Service Providers (CSPs) and non-expert investigator with each accounting for 9.02% of the participants as shown in Figure 21 above. Additionally, the eighth important factor was accessibility to data and this accounted for 8.20% of the participants as shown in Figure 21 above. Also, the ninth important factor had two factors which are spread of data in the cloud and reliability and privacy which each accounted for 7.38% of the participants as shown in Figure 21 above. Last but not least, was the eleventh most important factor which was decline in expert knowledge and this accounted for 5.74% of the participants as shown in Figure 21 above. However, it should be noted that this ranking was based on the percentage of participants who deem that factor as the most important to their respective organisations. In view of this, the importance of different factors

may differ from organisation to organisation or from industry to industry which may be influenced by the size and environmental context in which the organisation operates.

4.5 Automation of cyber forensics investigation performance

This section discusses the automation of cyber forensics investigation performance. This was measured using three metrics which are cyber security incidents, number of successful prosecutions and productivity & operational costs. Furthermore, the results from these three metrics were used to compute the mean score which became the metric to measure the automation of cyber forensics investigation performance. However, in order to do so the Cronbach's alpha value for reliability were computed as well. These shall be discussed in detail in the following subsections.

4.5.1 Cyber security incidents

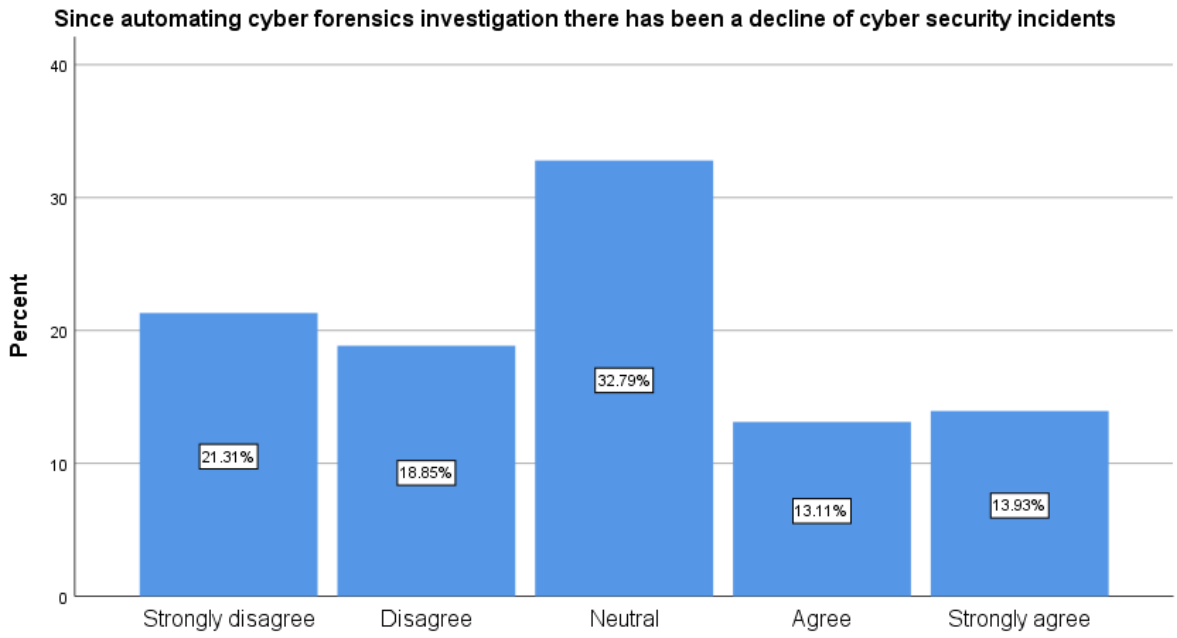

Figure 22: Cyber security incidents

The results in Figure 22 above indicate that most of the participants did not experience a decline of cyber security incidents since automation cyber forensics investigation. This was supported by a total disagreeableness score of 40.16% as shown in Figure 22 above. Furthermore, the other participants indicated that they had experienced a decline of cyber security incidents and this accounted for a total agreeableness score of 27.04%. However, 32.79% of the participants were neutral on this matter. These findings suggests that most of the participants were not experiencing a decline in cyber security incidents even after

automating the cyber forensics investigation process. This calls for further investigation to determine the root causes for this and potentially resolve them before serious consequences are experienced by these organisations. However, James and Gladyshev (2013) indicate that the automation of cyber forensics investigation is not an easy endeavour but is influenced by technological, political and social issues which may both facilitate or hamper the automation process. Accordingly, this may also have a bearing of the number of cyber security incidents that would be encountered by the organisations under consideration.

4.5.2 Number of successful prosecutions

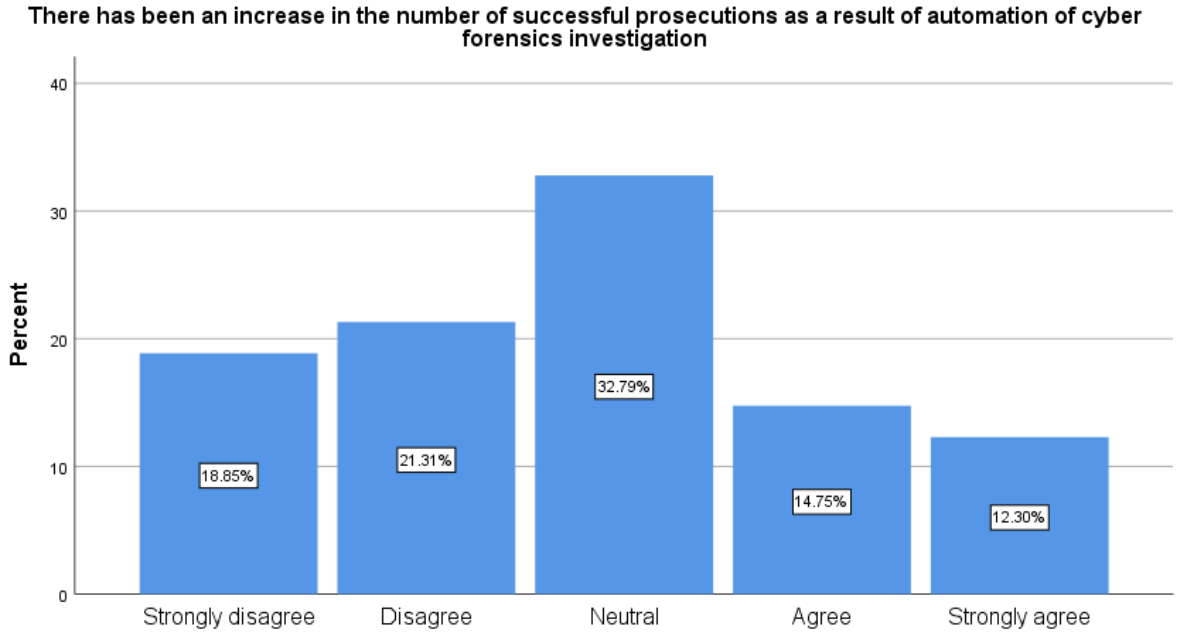

Figure 23: Number of successful prosecutions

Most of the participants in this study indicated that they did not experience any increase in the number of successful prosecutions as a result of automation of cyber forensics investigation. This accounted for a total disagreeableness score of 40.16% as shown in Figure 23 above. On the other hand, some of the participants indicated that they did experience an increase in the number of successful prosecutions and these accounted for a total agreeableness score of 27.05% as shown in Figure 23 above. Moreover, 32.79% of the participants indicated that there were neutral of this matter. It should be noted that successful prosecutions may take long to be realised as legal proceedings can typically take years in some cases. This could be the reasons why some of the participants took a neutral stance on this matter. According to Interpol (2023), digital forensics is useful in extracting data from electronic devices to obtain evidence that can be used for prosecution in the court of law. Although digital evidence is

integral for criminal investigation and prosecutions, it also has many challenges associated with it due to a number of reasons which include but not limited to: i) continuous advancements and changes in technology which the investigation needs to keep up with, ii) the need to communicate the technicalities associated with the evidence and iii) social political environment that leaves little room for error especially in data privacy (Miller, 2022). In view of this, it is reasonable to presume that even though the digital evidence might be there, it is still quite difficult (if not impossible) to maintain a 100% prosecution rate due to the evidence being disqualified or not sufficient to lead to a prosecution.

4.5.3 Productivity and operational costs

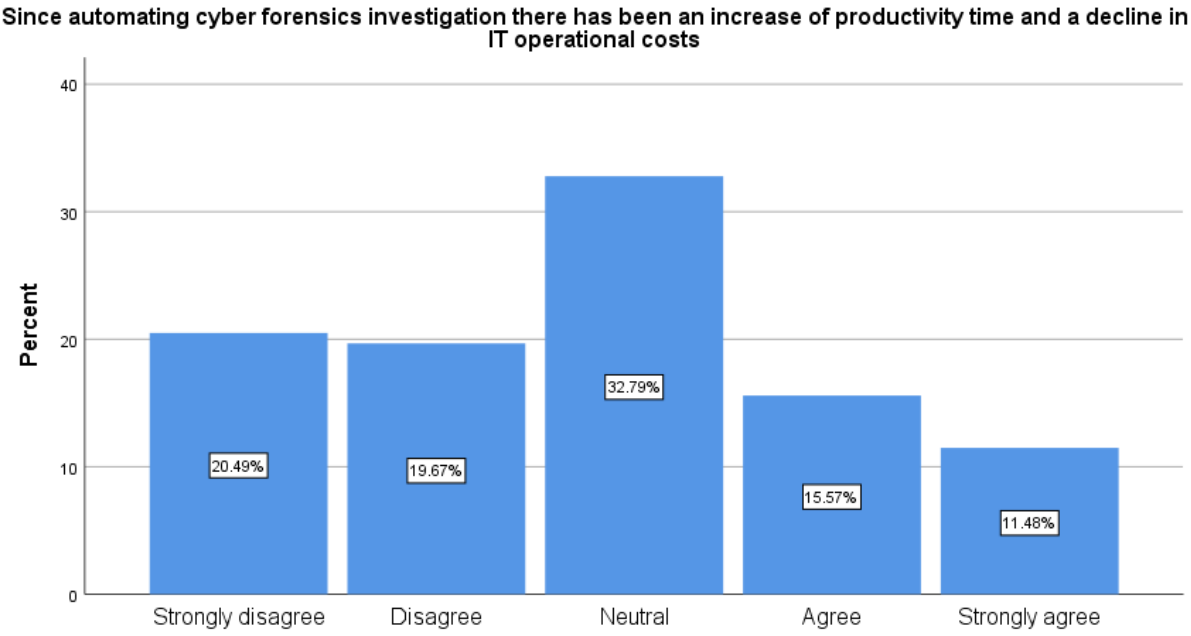

Figure 24: Productivity and operation costs

Most of the participants in the study indicated that they did not experience an increase of productivity time and a reduction in IT operational costs since automation cyber forensics investigation. This accounted for a total disagreeableness score of 40.16% of the participants as shown in Figure 24 above. However, some of the participants also indicated that when they automated cyber forensics investigation, they experience an increase in productivity time and a reduction in IT operational costs. This is supported by a total agreeableness score of 27.05% as shown in Figure 24 above. Nonetheless, 32.79% of the participants indicated that they were neutral in this matter. However, it is also important to note that the automation of cyber forensics investigation is expected to enhance the productivity and reduce operational costs due to improved operational efficiency (Avian, 2022). Despite this, it seems that a minority of

the participants were experiencing the better outcomes in terms of productivity and operation costs.

4.5.4 Reliability statistics

Table 4: Cronbach's alpha for automation of cyber forensics investigation performance

Reliability Statistics	
Cronbach's Alpha	N of Items
.965	3

Automation of cyber forensics investigation performance was computed using 3 questionnaire items that were combined. According to Arifin and Malaysia (2018), Cronbach's alpha values above 0.7 are acceptable. In view of this, the Cronbach's alpha value was computed to ensure reliability of the questionnaire instrument. The Cronbach's alpha value was 0.965 as shown in Table 4, which suggested that the questionnaire instrument used to measure the automation of cyber forensics investigation performance was reliable.

4.6 Correlations

The correlations between the independent and dependent variables in this study are shown in Table 5 below. However, the most important correlations in the study are between the dependent variable and the independent variables and these are highlighted in yellow in Table 5.

Table 5: Correlations of variables

Correlations													
		Automation of Cyber Forensics Investigation Performance	Accessibility to data	Unstable data	Dependency on Cloud Service Providers (CSPs)	Minimise time, maximise coverage	Decline in expert knowledge	Hardware requirements	Profiling and event construction	Artificial intelligence	Non-expert investigator	Spread of data in the cloud	Reliability and privacy
Automation of Cyber Forensics Investigation Performance	Pearson Correlation	1											
	Sig. (2-tailed)												
	N	122											
Accessibility to data	Pearson Correlation	.659**	1										
	Sig. (2-tailed)	.000											
	N	122	122										
Unstable data	Pearson Correlation	-.165	-.238**	1									
	Sig. (2-tailed)	.069	.008										
	N	122	122	122									
Dependency on Cloud Service Providers (CSPs)	Pearson Correlation	.658**	.332**	-.049	1								
	Sig. (2-tailed)	.000	.000	.595									
	N	122	122	122	122								
Minimise time, maximise coverage	Pearson Correlation	.344**	.196*	.044	.563**	1							
	Sig. (2-tailed)	.000	.030	.630	.000								
	N	122	122	122	122	122							
Decline in expert knowledge	Pearson Correlation	.311**	.247**	-.067	.409**	.631**	1						
	Sig. (2-tailed)	.000	.006	.464	.000	.000							
	N	122	122	122	122	122	122						
Hardware requirements	Pearson Correlation	-.321**	-.322**	.072	-.469**	-.642**	-.616**	1					
	Sig. (2-tailed)	.000	.000	.432	.000	.000	.000						
	N	122	122	122	122	122	122	122					
Profiling and event construction	Pearson Correlation	.194*	.091	-.033	.353**	.618**	.503**	-.483**	1				
	Sig. (2-tailed)	.032	.319	.722	.000	.000	.000	.000					
	N	122	122	122	122	122	122	122	122				
Artificial intelligence	Pearson Correlation	.077	.104	-.085	.079	.125	.063	.001	.136	1			
	Sig. (2-tailed)	.397	.253	.350	.388	.171	.489	.996	.135				
	N	122	122	122	122	122	122	122	122	122			
Non-expert investigator	Pearson Correlation	-.235**	-.105	-.031	-.333**	-.579**	-.487**	.325**	-.588**	-.107	1		
	Sig. (2-tailed)	.009	.250	.731	.000	.000	.000	.000	.000	.240			
	N	122	122	122	122	122	122	122	122	122	122		
Spread of data in the cloud	Pearson Correlation	-.135	-.160	-.013	-.202*	-.349**	-.518**	.346**	-.367**	-.044	.380**	1	
	Sig. (2-tailed)	.140	.079	.886	.025	.000	.000	.000	.000	.634	.000		
	N	122	122	122	122	122	122	122	122	122	122	122	
Reliability and privacy	Pearson Correlation	-.145	-.213*	.019	-.261**	-.469**	-.493**	.387**	-.325**	-.041	.308**	.397**	1
	Sig. (2-tailed)	.111	.019	.836	.004	.000	.000	.000	.000	.652	.001	.000	
	N	122	122	122	122	122	122	122	122	122	122	122	122

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

4.6.1 Hypotheses testing

This section presents the findings on the hypotheses that were generated in the literature review chapter, and these are shown in Table 6 below.

Table 6: Results for hypotheses testing

Hypothesis	r	Sig. level
H_{1a}: *Accessibility to data facilitates automation of cyber forensics investigation performance during the preparation and identification phase.*	0.659	Yes, 0.01
$H_{1b:}$ *Unstable data impedes the automation of cyber forensics investigation performance during the preparation and identification phase*	-0.165	No
H_{1c}: *The use of Trust Cloud facilitates the automation of cyber forensics investigation performance during the collection and preservation phase.*	0.658	Yes, 0.01
H_{1d}: *DFaaS facilitates the automation of cyber forensics investigation performance during the collection and preservation phase.*	0.344	Yes, 0.01
H_{1e}: *PBF facilitates the automation of cyber forensics investigation performance during the examination and analysis phase.*	0.311	Yes, 0.01
H_{1f}: *Hardware requirements impedes the automation of cyber forensics investigation performance during the examination and analysis phase.*	-0.321	Yes, 0.01
H_{1g}: *Profiling and event reconstruction facilitate the automation of cyber forensics investigation performance during the examination and analysis phase*.	0.194	Yes, 0.05
H_{1h}: *AI facilitates the automation of cyber forensics investigation performance during the examination and analysis phase*.	0.077	No
H_{1i}: *Non-expert investigator impedes the automation of cyber forensics investigation performance during the presentation phase.*	-0.235	Yes, 0.01
H_{1j}: *Spread of data in the cloud impedes the automation of cyber forensics investigation performance during the presentation phase.*	0.135	No
H_{1k}: *Reliability and privacy concerns impede the automation of cyber forensics investigation performance during the presentation phase.*	0.145	No
r – Pearson correlation coefficient; **Sig. level** – significant level		

4.7 Multiple regression analysis

The multiple regression analysis of the study is shown in Table 7 below. It was computed using six control variables (CV_1 – CV_6) which are age, education, employment status, province, job title, tenure and industry. Moreover, eleven independent variables represented by IV_1 – IV_{11} were accessibility to data, unstable data, dependency on CSPs, minimise time, maximise coverage, decline in expert knowledge, hardware requirements, profiling and event construction, artificial intelligence, non-expert investigator, spread of data in the cloud and reliability and privacy as shown in Table 7 below. Furthermore, when computing the multiple regression analysis; Model 1 consisted of the control variables whilst Model 2 consisted of the independent variables. However, it should be noted that Model 2 independent variables were computed individually and consolidated to form Table 7. The IBM SPSS Windows version 26 has a limitation of eight model blocks. In view of this, for the ninth to the eleventh independent

variables; Block 1 were the control variables, Block 2 were independent variables IV_1 – IV_8, Block 3 IV_9, Block 4 IV_{10} and Block 5 IV_{11}.

Also, Model 1 shows the analysis of the linear relationship between control variables and the dependent variable (DV). This was represented by the following equation:

$$DV = \beta_0 + \beta_1 CV_1 + \beta_2 CV_2 + \beta_3 CV_3 + \beta_4 CV_4 + \beta_5 CV_5 + \beta_6 CV_6$$

On the other hand, Model 2 shows the analysis of the linear relationship between the dependent variable and the independent variables. This was represented by the following equation:

$$DV = Model\ 1 + \beta_7 IV_1 + \beta_8 IV_2 + \beta_9 IV_3 + \beta_{10} IV_4 + \beta_{11} IV_5 + \beta_{13} IV_6 + \beta_{14} IV_7 + \beta_{15} IV_8 + \beta_{16} IV_9 + \beta_{17} CV_{10} + \beta_{18} CV_{11}$$

Nevertheless, the data in Table 6 below seem to suggest that Model 1 did not show a good model fit between the dependent variable and the control variables. Furthermore, it appears as if Model 2 was a good fit as majority of the independent variables had a significant relationship with the dependent variables except for artificial intelligence and reliability & privacy.

The results in Table 7 and Table 8 indicate that accessibility to data and dependency on CSPs were significant at a 0.01 level of influencing the automation of cyber forensics investigation performance of organisations. This accounted for a 37.8% and 17.5% change in the cyber forensics investigation performance whilst 10.8% was influenced by the control variables. In light of this, IV_2: Unstable data, IV_4: Minimise time, maximise coverage, IV_5: Decline in expert knowledge, IV_6: Hardware requirements, IV_7: Profiling and event construction, IV_8: Artificial intelligence, IV_9: Non-expert investigator, IV_{10}: Spread of data in the cloud and IV_{11}: Reliability and privacy accounted for 0.0%, 0.1%, 0.0%, 0.6%, 0.0%, 0.1%, 0.2%, 0.1% and 0.6% respectively as shown in Table 4 below. These findings indicate that these independent variables had no significant influence or impact on the cyber forensics investigation performance.

Table 7: Model summary

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.329[a]	.108	.053	1.19944
2	.823[b]	.678	.622	.75799
a. Predictors: (Constant), Industry, Education, Employment Status, Job Title, Tenure, Age, Province				
b. Predictors: (Constant), Industry, Education, Employment Status, Job Title, Tenure, Age, Province, Hardware requirements, Artificial intelligence, Unstable data, Non-expert investigator, Accessibility to data, Reliability and privacy, Spread of data in the cloud, Dependency on Cloud Service Providers (CSPs), Profiling and event construction, Decline in expert knowledge, Minimise time, maximise coverage				

Furthermore, it should be noted that the R square (R^2) change attributed to the control variables was 10.8% and when combined with the independent variables produced a R^2 change of 67.8% as shown in Table 6 above. This effectively means that the independent variables were responsible for 58.0% of the change experienced by the dependent variable. Moreover, considering that the other independent variables except accessibility to data and dependency on CSPs had no significant influence on the dependent variable as these two were responsible for 55.3% variation of the dependent variable. These findings suggest that accessibility to data and dependency on CSPs are the two most important factors that affect the automation of cyber forensics investigation based on the data collected in this study. In view of this, cyber forensics investigation performance can be significantly improved by enhancing accessibility to data and reducing dependency on CSPs.

Table 8: Regression analysis

REGRESSION ANALYSIS												
	Model 1	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**	**Model 2**
Constant	5.420	1.884	1.886	.359	.442	.400	-.270	-.193	5.172	5.842	5.824	5.979
CV_1: Age	-.243	-.064	-.064	-.028	-.032	-.030	-.024	-.025	-.239	-.222	-.206	-.202
CV_2: Education	-.159	-.068	-.068	-.025	-.021	-.021	-.021	-.019	-.155	-.175	-.167	-.185
CV_3: Employment status	-.384	-.172	-.172	-.045	-.039	-.044	-.042	-.044	-.376	-.362	-.369	-.354
CV_4: Province	-.062	-.070	-.070	-.033	-.034	-.034	-.031	-.033	-.063	-.045	-.047	-.050
CV_5: Job title	-.053	-.044	-.044	-.028	-.029	-.028	-.030	-.030	-.056	-.075	-.064	-.056
CV_6: Tenure	.101	.115	.115	.039	.039	.041	.038	.036	.099	.099	.105	.100
CV_7: Industry	-.042	-.023	-.023	-.013	-.012	-.012	-.015	-.015	-.042	-.048	-.050	-.050
IV_1: Accessibility to data		.629**										
IV_2: Unstable data			-.001									
IV_3: Dependency on CSPs				.419**								
IV_4: Minimise time, maximise coverage					-.049							
IV_5: Decline in expert knowledge						.020						
IV_6: Hardware requirements							.136					
IV_7: Profiling and event construction								-.027				
IV_8: Artificial intelligence									.050			
IV_9: Non-expert investigator										-.228		
IV_{10}: Spread of data in the cloud											-.069	
IV_{11}: Reliability and privacy												-.094
R^2	10.8%	48.6%	48.6%	66.1%	66.2%	66.2%	66.8%	66.8%	66.9%	67.1%	67.3%	67.9%
R^2 change	10.8%	37.8%	0.0%	17.5%	0.1%	0.0%	0.6%	0.0%	0.1%	0.2%	0.1%	0.6%
F-value	1.971	13.348	11.760	21.643	19.614	17.827	16.695	15.382	1.729	2.196	1.998	1.863
F-value change	1.971	83.059**	.000	57.347**	.430	.046	1.711	.106	.144	.520	.471	2.077
VIF range	1.035 – 1.070	1.042 – 1.107	1.045 – 1.149	1.050 – 1.233	1.054 – 1.701	1.055 – 2.119	1.055 – 2.439	1.058 – 2.814	1.067 – 2.844	1.070 – 3.036	1.081 – 3.053	1.122 – 3.172

*: p<0.05; **:p<0.01; Dependent variable: Automation of cyber forensics investigation performance

Table 9: Coefficients

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	5.420	1.021		5.310	.000
	Age	-.243	.139	-.160	-1.745	.084
	Education	-.159	.138	-.104	-1.156	.250
	Employment Status	-.384	.221	-.157	-1.736	.085
	Province	-.062	.041	-.139	-1.518	.132
	Job Title	-.053	.096	-.050	-.555	.580
	Tenure	.101	.083	.110	1.218	.226
	Industry	-.042	.038	-.102	-1.124	.263
2	(Constant)	-.376	1.313		-.287	.775
	Age	-.039	.092	-.026	-.427	.670
	Education	-.007	.091	-.005	-.079	.937
	Employment Status	-.065	.147	-.026	-.441	.660
	Province	-.026	.028	-.057	-.923	.358
	Job Title	-.051	.064	-.048	-.797	.428
	Tenure	.038	.056	.041	.684	.496
	Industry	-.013	.025	-.030	-.506	.614
	Accessibility to data	.510	.065	.517	7.885	.000
	Unstable data	-.037	.127	-.018	-.289	.773
	Dependency on Cloud Service Providers (CSPs)	.444	.066	.493	6.679	.000
	Minimise time, maximise coverage	.008	.109	.007	.070	.945
	Decline in expert knowledge	.099	.109	.080	.902	.369
	Hardware requirements	.121	.109	.094	1.110	.270
	Profiling and event construction	-.040	.090	-.036	-.443	.658
	Artificial intelligence	-.034	.087	-.023	-.387	.699
	Non-expert investigator	-.068	.086	-.062	-.788	.432
	Spread of data in the cloud	.036	.086	.030	.416	.678
	Reliability and privacy	.124	.086	.101	1.441	.153
a. Dependent Variable: Automation of Cyber Forensics Investigation Performance						

4.8 Chapter summary

Most of the participants were aged 31 to 40 years, had a diploma as their highest level of education, had a job title of security specialist or professional within their organisation and have been working for 6 to 9 years within their current organisation. The participants were equally distributed between those who were employed full time and those who were working part time. Furthermore, most of the participants came from the North-West province but other

provinces were also fairly represented. Additionally, most of the participants were working in the finance industry but other industries were also decently represented. Besides all the participants indicated that there were automating the cyber forensics investigation within their organisation. The results also indicated that accessibility to data, dependency on CSPs, minimise time, maximise coverage, decline in the expert knowledge, profiling and event construction and artificial intelligence were factors that most of the organisations represented in this study were performing well on. However, dealing with unstable data, hardware requirements, non-expert investigator, spread of data in the cloud and reliability and privacy proved to be a challenge for many of the organisations that were represented in this study. Additionally, the findings also revealed that accessibility to data and dependency on CSPs were the two most significant factors that affected the automation of cyber forensics investigation performance. The next chapter will present the conclusions and recommendations of the study.

CHAPTER 5: CONCLUSIONS & RECOMMENDATIONS

5.1 Introduction

Cyber forensics is a type of forensics science and is known as the identifying, collecting, preserving, analysing and presenting of digital evidence. For these processes or phases to work efficiently, a control is put in place called the Chain of Custody, which regulates the management of evidence from start to finish. There was a need to understand the process and in turn, it needs to be efficient enough to keep up with technological advances, reduce costs and avoid unnecessarily labour-intensive procedures. The researcher's intention was to identify the factors that will influence automation in the cyber forensics phases as it will improve industry knowledge and ability to identify barriers and inefficiencies to the process. The objective of this chapter is to focus on what these results imply and then finally to draw conclusions, with an additional section for recommendations for future studies surrounding the automation of the cyber forensics process. As a result, this chapter shall be divided into two main sections which are conclusions and recommendations.

5.2 Conclusions

This section will be divided into two subsections which are conclusions on hypotheses testing and conclusions on the objectives of the study.

5.2.1 Conclusions on hypotheses testing.

The study generated eleven hypotheses from the literature review and these were stated during the data analysis process in the previous study. The conclusions on the hypotheses testing are shown in Table 9 below:

Table 10: Conclusions on hypotheses testing.

Hypothesis	r	Sig. level	Conclusion
H_{1a}: *Accessibility to data facilitates automation of cyber forensics investigation during the preparation and identification phase.*	0.659	Yes, 0.01	It was concluded that accessibility to data facilitates automation of cyber forensics investigation during the preparation and identification phase and this was significant at 0.01 level.
$H_{1b:}$ *Unstable data impedes the automation of cyber forensics investigation during the preparation and identification phase*	-0.165	No	It was concluded that unstable data impedes the automation of cyber forensics investigation during the preparation and identification phase but is was not significant at 0.05 level.
H_{1c}: *The use of Trust Cloud facilitates the automation of cyber forensics investigation during the collection and preservation phase.*	0.658	Yes, 0.01	It was concluded that the use of Trust Cloud facilitates the automation of cyber forensics investigation during the collection and presentation phase and this was significant at 0.01 level.
H_{1d}: *DFaaS facilitates the automation of cyber forensics investigation during the collection and preservation phase.*	0.344	Yes, 0.01	It was concluded that DFaaS facilitates the automation of cyber forensics investigation during the collection and preservation phase and this was significant at 0.01 level.
H_{1e}: *PBF facilitates the automation of cyber forensics investigation during the examination and analysis phase.*	0.311	Yes, 0.01	It was concluded that PBF facilitates the automation of cyber forensics investigation during the examination and analysis phase and this was significant at 0.01 level
H_{1f}: *Hardware requirements impedes the automation of cyber forensics investigation during the examination and analysis phase.*	-0.321	Yes, 0.01	It was concluded that hardware requirements impedes that automation of cyber forensics investigation during the examination and analysis phase and this was significant at 0.01 level.
H_{1g}: *Profiling and event reconstruction facilitates the automation of cyber forensics investigation during the examination and analysis phase*.	0.194	Yes, 0.05	It was concluded that profiling and event reconstruction facilitates the automation of cyber forensics investigation during the examination and analysis phase and this was significant at 0.05 level.
H_{1h}: *AI facilitates the automation of cyber forensics investigation during the examination and analysis phase*.	0.077	No	It was concluded that AI facilitates the automation of cyber forensics investigation during the examination and analysis phase but is was not significant at 0.05 level.
H_{1i}: *Non-expert investigator impedes the automation of cyber forensics investigation during the presentation phase.*	-0.235	Yes, 0.01	It was concluded that non-expert knowledge investigator impedes the automation of cyber forensics investigation during the presentation phase and this was significant at 0.01 level.
H_{1j}: *Spread of data in the cloud impedes the automation of cyber forensics investigation during the presentation phase.*	0.135	No	It was concluded that spread of data in the cloud impedes the automation of cyber forensics investigation during the presentation phase but is was not significant at 0.05 level.
H_{1k}: *Reliability and privacy concerns impedes the automation of cyber forensics investigation during the presentation phase.*	0.145	No	It was concluded that reliability and privacy concerns impedes the automation of cyber forensics investigation during the presentation phase but is was not significant at 0.05 level.
r – Pearson correlation coefficient; **Sig. level** – significant level			

5.2.2 Conclusions on the objectives of the study

The objectives of the study were as follows:

- To identify factors that affect the automation of cyber forensics investigation,
- To determine the extent to which the factors influence the automation of cyber forensics investigation performance.

Objective 1: To identify factors that affect the automation of cyber forensics investigation.

Eleven factors that affect the automation of cyber forensics investigation identified in this study were the following:

- Accessibility to data – It was revealed that in cases were data is not recorded independently, cyber forensics experts are able to assess whether an investigation is needed or not. But it becomes a greater challenge to deal with data in different information centres around the world. In view of this, human intervention would also be required at some point in the process and this creates challenges for the automation of cyber forensics investigation. (Morales-Ferreira *et al.*, 2018; Rane and Dixit, 2019)
- Unstable data – It is known that in the event that a gadget is switched off, automation of cyber forensics investigation cannot be done due to inability to utilize system tools to automatically identify the evidence on the gadget. This then requires a human intervention to switch on the gadget and relink misplaced registry and brief web records which can be done remotely or physically (MacDermott *et al.*, 2018).
- Dependency on Cloud Service Providers (CSPs) – Due to data protection and privacy policies CSPs are typically strict and unwilling to share client's data and this creates problems for automating the cyber forensics investigation process. Additionally, CSPs usually keep data for a short period of time as it is either overwritten or deleted and could also be distributed across undisclosed locations which further impedes the automation of cyber forensics investigation. The cloud systems are also not designed with a focus on cyber forensics investigation but with a focus on operational use. In view of this, CSPs typically prioritise restoring the cloud system online for operational use without prioritising the integrity of the evidence (Alqahtany, 2017; Morales-Ferreira *et al.*, 2018).
- Minimise time, maximise coverage – DFaaS is a cloud-based service that allows forensic copies to be made of devices and stored to a central storage that is used to extract or collect evidence automatically for analysis, making a significant contribution to automation of the collection phase. Forensic copies are made from various devices and platforms; thus, a huge area of data is covered. The accessibility of forensic copies

ensures that evidence can be collected when the need arises, saving time and resources (Du *et al.*, 2017).

- Decline in expert knowledge – Examination and investigation in cyber forensics is not usually fully automated but it has a certain level of computerization that is utilised in 'push-button forensics' (PBF). These PBF apparatuses permit scientific examiners to perform complicated investigation capacities, simply by knowing which buttons to press. This may create a stagnation impact on the information learning capacity of cyber forensics experts as they can initiate the investigation by just pushing a button. In view of this, the higher level of mechanization PBF devices can impact the quality of the investigation as the cyber forensics expert may be unable to prove how the evidence was gathered in the process unveiling the lack of expertise in the cyber forensics expert which may negatively affect the successful prosecution in the court of law (Horsman, 2020; Vallor, 2017).
- Hardware requirements – High performance machines ought to be utilised for the automatization of the cyber forensics investigation because it syphers through enormous datasets. In view of this, automation of cyber forensics investigation requires computers with higher performance specifications and a lack thereof may impede the successfulness of the process as the time factor is an important component in this regard (Atlam *et al.*, 2020).
- Profiling and event construction – Computerized forensics gives little hypothetical premise to back and connect discoveries of an investigation. In view of this, profiling and event construction help to validate and evaluate the investigation process. As such, there is a need to automate ad-hoc validation activities to circumvent bias and preconceptions (Al Mutawa *et al.*, 2019; Horsman, 2020).
- Artificial intelligence – Intelligent forensics comprises of an assortment of instruments and procedures from AI and machine learning. Modelling and social network analysis helps to cut down on the amount of time spent on searching for digital proof (Wylot *et al.*, 2018).
- Non-expert investigator – When a non-expert investigator presents evidence that was automatically discovered, it can raise a concern for lawyers and judges. It is important that an expert witness validates all the forensics software and tools that have been used during the investigation. If this fails to happen, evidence will lose credibility and become inadmissible in court.
- Spread of data in the cloud - Due to the vast infrastructure of the cloud, expert witnesses find it challenging to prove in which country an incident has occurred. This in turn, creates misalignment with the investigators by establishing which laws to obey

based on the specific country. Therefore, investigators must explain the technicalities of how this phase was automated to acquire the evidence from the cloud environment. In turn, the jury find it hard to comprehend complexities of cloud forensics (Simou *et al.*, 2016).

- Reliability and privacy - During this phase of automation, reliability and privacy is crucial for the law. The concept of reliability is connected to the accuracy of the forensic tool (Hughes and Karabiyik, 2020). The safe keeping and security of data will always be vulnerable. However, to mitigate this, tools need to drastically improve, and the cyber forensic process should become less cumbersome and more automated (Hughes and Karabiyik, 2020).

Objective 2: To determine the extent to which the factors influence the automation of cyber forensics investigation performance.

The multiple regression analysis that was conducted on the collected revealed that Accessibility to data and Dependency on CSPs (i.e., use of Trust Cloud facilities) were the two most important factors that influenced the Automation of cyber forensics investigation performance. It was revealed that accessibility to data was responsible for 37.8% of the variation in the Automation of cyber forensics investigation performance whilst Dependency of CSPs was responsible for 17.5% of the variation. Furthermore, the influence of these two factors were significant at 0.01 level. However, the other 9 factors were not significant at 0.05 level, suggesting that they did not have much influence on the Automation of cyber forensics investigation performance. In light of this, these two independent variables had a significant impact on the automation of cyber forensics investigation performance and because of this, they can be deemed as the most important factors that influence the automation of cyber forensics investigation performance.

It is important to note that the success or failure of the automation of cyber forensics investigation industry relies on the accessibility to data. If data is inaccessible cyber forensics investigation cannot be done at all, let alone its automation. In this regard, accessibility to data can be adversely affected by data stored in clouds that are distributed over several locations globally (Morales-Ferreira *et al.*, 2018). As such, human intervention may be necessary to authorise and manually access the data. In light of this, ensuring that data is accessible will enhance the effectiveness of the automation of the cyber forensics investigation.

On the other hand, even if data is accessible, it must be accessed in a lawful manner. In view of this, CSPs are not always forthcoming with their data as they are required by the law to keep the data private and confidential. Besides, the automation of the cyber forensics

investigation process cannot be done on data that is protected by privacy and confidentiality laws and that could also be distributed across the globe at undisclosed locations may also hinder the automation of cyber forensics investigation process (Morales-Ferreira *et al.*, 2018). Therefore, the use of Trust Cloud facilities is a necessary step to ensure that cyber forensics investigations are automated.

In light of the discussion above, it is evident why these two independent variables were regarded or found to be the most important factors that influence the Automation of cyber forensics investigation performance.

5.3 Recommendations

Future research in data accessibility could help forensics investigators in executing their tasks more efficiently. Constant awareness and continuous research of forensic readiness will add great value to the automation of the cyber forensics process. We are living in a world where technology surpasses our understanding, so are cyber criminals. Therefore, further research of trust and reliability amongst employees can also be a recommendation to add to the body knowledge, as most of the times, it's employees or colleagues that is directly or indirectly the cause of data breaches.

5.4 Chapter summary

Eleven factors that affect cyber forensics investigation processes were identified in this study from the literature and these were: 1) Accessibility to data, 2) Unstable data, 3) Dependency on Cloud Service Providers (CSPs), 4) Minimise time, maximise coverage, 5) Decline in expert knowledge, 6) Hardware requirements, 7) Profiling and event construction, 8) Artificial intelligence, 9) Non-expert investigator, 10) Spread of data in the cloud and 11) Reliability and privacy. Moreover, data was collected from cyber forensics and IT security experts within the republic of South Africa and this data was statistically analysed using descriptive statistics, correlations, regression analysis and hypothesis testing. Moreover, correlations between these 11 independent variables were conducted against the dependent variable (Automation of cyber forensics investigation performance). Out of the 11 independent variables, only 7 variables (Accessibility to data, The use of Trust Cloud, DFaaS, PBF, Hardware requirements, Profiling and event reconstruction, Non-expert investigator) were found to be statistically significant at either 0.01 or 0.05 level. However, a multiple regression analysis revealed that Accessibility to data and Dependency on CSPs independent variables had significant positive influenced on the Automation of cyber forensics investigation performance. These accounted for 37.8% and 17.5% of change in Automation of cyber forensics investigation performance

respectively. As such, this result suggest that more resources should be focused on Accessibility to data and Dependency on CSPs to optimise and enhance the Automation of cyber forensics investigation performance.

REFERENCES

Aamir, M. PLC for optimization and reliability in the field of Automation. 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), 2017. IEEE, 1-1.

Acemoglu, D. & Restrepo, P. 2018. Artificial intelligence, automation, and work. *The economics of artificial intelligence: An agenda.* University of Chicago Press.

Achari, P. D. 2014. *Research Methodology: A guide to ongoing research scholars in management*, 1st ed., Asia: Horizon Books.

Ademu, I. O. 2013. *A Comprehensive Digital Forensic Investigation Model and Guidelines for Establishing Admissible Digital Evidence.* PhD thesis, University of East London.

Agarwal, A., Gupta, M., Gupta, S. & Gupta, S. C. 2011. Systematic digital forensic investigation model. *International Journal of Computer Science and Security (IJCSS),* 5(1)**,** pp 118-131.

Ahuett-Garza, H. & Kurfess, T. 2018. A brief discussion on the trends of habilitating technologies for Industry 4.0 and Smart manufacturing. *Manufacturing Letters,* 15**,** pp 60-63.

Al Fahdi, M. 2016. *Automated Digital Forensics and Computer Crime Profiling.* PhD thesis, University of Plymouth.

Al Mutawa, N., Bryce, J., Franqueira, V. N., Marrington, A. & Read, J. C. 2019. Behavioural digital forensics model: Embedding behavioural evidence analysis into the investigation of digital crimes. *Digital Investigation,* 28**,** pp 70-82.

Al-khateeb, S. & Agarwal, N. 2019. Social Cyber Forensics (SCF): Uncovering Hidden Relationships. *Deviance in Social Media and Social Cyber Forensics.* Springer.

Alawadhi, I., Read, J. C., Marrington, A. & Franqueira, V. N. 2015. Factors influencing digital forensic investigations: Empirical evaluation of 12 years of dubai police cases. *Journal of Digital Forensics, Security and Law,* 10(4)**,** pp 7-16.

Alawadhi, I. M. 2019. *Methods and factors affecting digital forensic case management, allocation and completion.* PhD thesis, University of Central Lancashire.

Alqahtany, S. 2017. *A forensically-enabled IaaS cloud computing architecture.* Doctorate degree, University of Plymouth.

Amit, K. 2018. *Artificial Intelligence and Soft Computing: Behavioral and cognitive modeling of the human brain*, 1st ed., Boca Raton: CRC Press.

Arafat, M. Y., Mondal, B. & Rani, S. 2017. Technical challenges of cloud forensics and suggested solutions. *Int. J. Sci. Eng. Res,* 8(8)**,** pp 1142-1149.

Arifin, W. N. & Malaysia, S. 2018. Calculating the Cronbach's alpha coefficients for measurement scales with" not applicable" option. *Universiti Sains Malaysia***,** pp 1-8.

Arnanda, E. 2020. *Political Islam throughout history of Indonesia: An analysis of Indonesian counterterrorism.* MA in Global Crime and Justice, The University of York.

Arshad, H., Abdullah, S., Alawida, M., Alabdulatif, A., Abiodun, O. I. & Riaz, O. 2022. A Multi-Layer Semantic Approach for Digital Forensics Automation for Online Social Networks. *Sensors,* 22(3)**,** pp 1115.

Atlam, H. F., Alenezi, A., Alassafi, M. O., Alshdadi, A. A. & Wills, G. B. 2020. Security, Cybercrime and Digital Forensics for IoT. *In:* Peng, S.-L., Pal, S. & Huang, L. (eds.) *Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm.* Cham: Springer International Publishing.

Avian. 2022. *Why it's worth it? 5 reasons for implementing automation for your digital forensics or e-discovery workflows* [Online]. Available: https://www.avian.dk/blog/5-reasons-for-

implementing-automation-for-your-digital-forensics-or-e-discovery-workflows#:~:text=Minimized%20errors,errors%20when%20accelerating%20case%20throughput. [Accessed 27 April 2023].

Banerjee, A., Chitnis, U., Jadhav, S., Bhawalkar, J. & Chaudhury, S. 2009. Hypothesis testing, type I and type II errors. *Industrial psychiatry journal,* 18(2)**,** pp 127.

Baryamureeba, V. & Tushabe, F. 2004. The enhanced digital investigation process model. Available: https://dfrws.org/wp-content/uploads/2019/06/2004_USA_paper-the_enhanced_digital_investigation_process_model.pdf [Accessed 30 July 2022].

Bhandayker, Y. R. 2019. A Study on the Research Challenges and Trends of Cloud Computing. *Research Review International Journal of Multidisciplinary,* 4(2)**,** pp 441-447.

Billings, C. E. 2018. *Aviation automation: The search for a human-centered approach*, 1st ed., Mahwah, NJ: CRC Press.

Boutell, M. & Luo, J. 2005. Beyond pixels: Exploiting camera metadata for photo classification. *Pattern recognition,* 38(6)**,** pp 935-946.

Chhabra, G. S., Singh, V. P. & Singh, M. 2020. Cyber forensics framework for big data analytics in IoT environment using machine learning. *Multimedia Tools and Applications,* 79(23)**,** pp 15881-15900.

Cho, J. Y. & Lee, E.-H. 2014. Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *Qualitative report,* 19(32)**,** pp 1-20.

Choo, K.-K. R. 2017. Contemporary digital forensic investigations. *Resource Material Series No. 97* [Online]. Available: https://unafei.or.jp/publications/pdf/RS_No97/No97_Australia_2.pdf [Accessed 3 July 2022].

Conlan, K., Baggili, I. & Breitinger, F. 2016. Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy. *Digital investigation,* 18**,** pp S66-S75.

Coppini, A. & Saliba, M. A. 2017. Towards practical guidelines for conversion from a fixed to a reconfigurable manufacturing automation system. *Procedia manufacturing,* 11**,** pp 1102-1111.

Creswell, J. W. & Creswell, J. D. 2018. *Research design: Qualitative, quantitative and mixed methods approaches*, 5th ed., Los Angeles, USA: Sage Publications Inc.

Damshenas, M., Dehghantanha, A., Mahmoud, R. & Bin Shamsuddin, S. Forensics investigation challenges in cloud computing environments. Proceedings Title: 2012 International Conference on Cyber Security, Cyber Warfare and Digital Forensic (CyberSec), 2012. IEEE, 190-194.

Das, S., Roy, K. & Nampi, T. 2020. Manufacturing, Control, and Automation. *Handbook of Research on Developments and Trends in Industrial and Materials Engineering.* IGI Global.

David, F. N. 2017. *Forces of production: A social history of industrial automation*, 1st ed.: Routledge.

De Mauro, A., Greco, M. & Grimaldi, M. 2016. A formal definition of Big Data based on its essential features. *Library Review,* 65(3)**,** pp 122-135.

Du, X., Le-Khac, N.-A. & Scanlon, M. 2017. Evaluation of digital forensic process models with respect to digital forensics as a service. *arXiv preprint arXiv:1708.01730***,** pp 1-10.

Emmanuel, N. 2019. *Analysis of Saunders research onion* [Online]. Available: https://thesismind.com/analysis-of-saunders-research-onion/ [Accessed 10 May 2020].

Endsley, M. R. 2018. Automation and situation awareness. *Automation and human performance: Theory and applications.* CRC Press.

Fisher, C. & Buglear, J. 2010. *Researching and writing a dissertation: An essential guide for business students*, ed.: Pearson Education.

Garfinkel, S. L. 2010. Digital forensics research: The next 10 years. *Digital Investigation,* 7**,** pp S64-S73.

Garfinkel, S. L., Parker-Wood, A., Huynh, D. & Migletz, J. 2010. An automated solution to the multiuser carved data ascription problem. *IEEE Transactions on Information Forensics and Security,* 5(4)**,** pp 868-882.

Gerda Síochána Inspectorate. 2015. Changing policing in Ireland. Available: https://www.gsinsp.ie/wp-content/uploads/2020/01/Changing-Policing-in-Ireland.pdf [Accessed 15 April 2023].

Glen, S. 2022. *Correlation coefficient: Simple definition, formula, easy steps* [Online]. Available: https://www.statisticshowto.com/probability-and-statistics/correlation-coefficient-formula/ [Accessed 6 July 2022].

Glisson, W. & Choo, K.-K. R. Introduction to the Minitrack on Cyber-of-Things: Cyber Crimes, Cyber Security and Cyber Forensics. Proceedings of the 51st Hawaii International Conference on System Sciences, 2018.

Goss, J. & Gladyshev, P. 2010. *Forensic triage: Managing the risk.* Master's thesis, University College Dublin.

Groover, M. P. 2016. *Automation, production systems, and computer-integrated manufacturing*, 1st ed.: Pearson Education India.

Gupta, B. 2016. *Interview questions in business analytics*, 1st ed., Berkeley, CA: Apress.

Hafner, K. & Markoff, J. 1991. *Cyberpunk: Outlaws and hackers on the computer frontier* 1st ed., New York: Touchstone.

Hayes, D. & Kyobe, M. The adoption of automation in cyber forensics. 2020 Conference on Information Communications Technology and Society (ICTAS), 2020. IEEE, 1-6.

Heale, R. & Twycross, A. 2015. Validity and reliability in quantitative studies. *Evidence-based nursing,* 18(3)**,** pp 66-67.

Hilburn, B. 2017. Dynamic decision aiding: the impact of adaptive automation on mental workload. *Engineering psychology and cognitive ergonomics.* Routledge.

Hitchcock, B., Le-Khac, N.-A. & Scanlon, M. 2016. Tiered forensic methodology model for Digital Field Triage by non-digital evidence specialists. *Digital investigation,* 16**,** pp S75-S85.

Homem, I. 2016. *Towards automation in digital investigations: Seeking efficiency in digital forensics in mobile and cloud environments.* PhD thesis, Department of Computer and Systems Sciences, Stockholm University.

Homem, I. 2018a. *Advancing automation in digital forensic investigations.* PhD thesis, Department of Computer and Systems Sciences, Stockholm University.

Homem, I. 2018b. *Advancing automation in digital forensic investigations.* Doctorate degree, Stockholm University.

Horsman, G. 2020. Opinion: Does the field of digital forensics have a consistency problem? *Forensic Science International: Digital Investigation,* 33**,** pp 300970.

Hughes, N. & Karabiyik, U. 2020. Towards reliable digital forensics investigations through measurement science. *Wiley Interdisciplinary Reviews: Forensic Science,* 2(4)**,** pp e1367.

Husain, M. S. & Khan, M. Z. 2019. *Critical Concepts, Standards, and Techniques in Cyber Forensics*, 1st ed., Hershey, PA, USA: IGI Global.

Interpol. 2023. *Digital forensics* [Online]. Available: https://www.interpol.int/en/How-we-work/Innovation/Digital-forensics [Accessed 29 April 2023].

Irons, A. & Lallie, H. S. 2014. Digital forensics to intelligent forensics. *Future Internet,* 6(3)**,** pp 584-596.

Israel, G. D. 2008. Determining sample size. Available: https://cdm.unfccc.int/filestorage/Z/T/P/ZTPTD6SMKP1ZVBM6LVNA3VJAAGLK8M/References.pdf?t=RTB8cWllbnRxfDBgaK6CSEd_ve-QVe8IS8YV [Accessed 18 October 2018].

Jaishankar, K. 2018. Cyber criminology as an academic discipline: history, contribution and impact. *International Journal of Cyber Criminology,* 12(1)**,** pp 1-8.

James, J. I. & Gladyshev, P. 2013. Challenges with Automation in Digital Forensic Investigations. Available: https://arxiv.org/pdf/1303.4498.pdf [Accessed 16 April 2023].

Jusas, V., Birvinskas, D. & Gahramanov, E. 2017. Methods and Tools of Digital Triage in Forensic Context: Survey and Future Directions. *Symmetry,* 9(4)**,** pp 49.

Kao, D.-Y. 2016. Cybercrime investigation countermeasure using created-accessed-modified model in cloud computing environments. *The Journal of Supercomputing,* 72(1)**,** pp 141-160.

Karafili, E., Cristani, M. & Viganò, L. A formal approach to analyzing cyber-forensics evidence. European Symposium on Research in Computer Security, 2018. Springer, 281-301.

Koleoso, R. A. 2018. *A Digital Forensics Investigation Model for Confidentiality, Integrity and Authenticity.* Doctorate degree, University of Lagos.

Kothari, S. & Hasija, H. Spiral Model for Digital Forensics Investigation. *In:* Thampi, S., Martínez Pérez, G., Westphall, C., Hu, J., Fan, C. & Gómez Mármol, F., eds. International Symposium on Security in Computing and Communication, 2017 Singapore. Springer, 312-324.

Krieger, N. 2012. Who and what is a “population”? Historical debates, current controversies, and implications for understanding “population health” and rectifying health inequities. *The Milbank Quarterly,* 90(4)**,** pp 634-681.

Kumar, R. 2019. *Programmable cyber networks for critical infrastructure.* Doctorate degree, University of Illinois at Urbana-Champaign.

Lienenlüke, L., Gründel, L., Storms, S., Herfs, W., Konigs, M. & Servos, M. Temporal and flexible automation of machine tools. 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), 2018. IEEE, 000335-000340.

Lillis, D., Becker, B., O'Sullivan, T. & Scanlon, M. 2016. Current challenges and future research areas for digital forensic investigation. Available: https://arxiv.org/pdf/1604.03850.pdf [Accessed 15 April 2023].

Lone, A. H. & Mir, R. N. 2019. Forensic-chain: Blockchain based digital forensics chain of custody with PoC in Hyperledger Composer. *Digital investigation,* 28**,** pp 44-55.

Lovato, M. 2017. *A race against the machine: has the threat of technological unemployment finally become true.* Master's thesis, Universita' Degli Studi Di Padova.

Lu, Y. & Da Xu, L. 2018. Internet of Things (IoT) cybersecurity research: A review of current research topics. *IEEE Internet of Things Journal,* 6(2)**,** pp 2103-2115.

Lutui, R. 2016. A multidisciplinary digital forensic investigation process model. *Business Horizons,* 59(6)**,** pp 593-604.

MacDermott, A., Baker, T. & Shi, Q. Iot forensics: Challenges for the ioa era. 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS), 2018. IEEE, 1-5.

Majaski, C., Khartit, K. & Kvilhaug, S. 2021. *What is hypothesis testing* [Online]. Available: https://www.investopedia.com/terms/h/hypothesistesting.asp [Accessed 6 July 2022].

Martinova, L. & Martinov, G. Automation of machine-building production according to industry 4.0. 2018 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC), 2018. IEEE, 1-4.

McKemmish, R. 1999. What is forensic computing? Available: https://www.aic.gov.au/sites/default/files/2020-05/tandi118.pdf [Accessed 3 July 2022].

Miller, C. M. 2022. A survey of prosecutors and investigators using digital evidence: A starting point. *Forensic Science International: Synergy***,** pp 100296.

Mody, S. & Nisbet, A. A centralised platform for digital forensic investigations in cloud-based environments. *In:* Valli, C., ed. The Proceedings of 15th Australian Digital Forensics Conference, 5-6 December 2017 Perth, Australia. Edith Cowan University.

Momoh, J. A. 2017. *Electric power distribution, automation, protection, and control*, 1st ed., Boca Raton: CRC press.

Moore, A. W., Anderson, B., Das, K. & Wong, W.-K. 2006. CHAPTER 15 - Combining multiple signals for biosurveillance. *In:* Moore, A. W., Anderson, B., Das, K. & Wong, W.-K. (eds.) *Handbook of biosurveillance.* Amsterdam: Academic Press.

Morales-Ferreira, P., Santiago-Duran, M., Gaytan-Diaz, C., Gonzalez-Compean, J., Sosa-Sosa, V. J. & Lopez-Arevalo, I. A data distribution service for cloud and containerized storage based on information dispersal. 2018 IEEE Symposium on Service-Oriented System Engineering (SOSE), 2018. IEEE, 86-95.

Moroz, E. 2018. Computer aided manufacturing processes using Lean Management and Lean Manufacturing methods. *Mechanik,* 91(7)**,** pp 535-537.

Naz, S. & Iraqi, K. M. 2019. Impact of Robotics Process Automation on Human Resources Management Practices in IT Sector of Pakistan. *International Journal of Business Studies,* 1(2)**,** pp.

Palmer, J., Llorens, B., Kaufman, S., Gibbons, C., Chowdhury, M., Chen, C. & Fu, X. Modeling Cyber Crimes and Investigations for Digital Forensics Education. Journal of The Colloquium for Information Systems Security Education, 2016. 23-23.

Polak, S. 2019. *Designing a system of automation of the process of banding furniture boards in an enterprise introducing a flexible production system.* Doctorate degree, Warsaw University of Technology.

Pollitt, M. A history of digital forensics. IFIP International Conference on Digital Forensics, 2010. Springer, 3-15.

Raghavan, S. 2013. Digital forensic research: current state of the art. *CSI Transactions on ICT,* 1**,** pp 91-114.

Rane, S. & Dixit, A. BlockSLaaS: Blockchain assisted secure logging-as-a-service for cloud forensics. International Conference on Security & Privacy, 2019. Springer, 77-88.

Richard III, G. G. & Roussev, V. 2006. Next-generation digital forensics. *Communications of the ACM,* 49(2)**,** pp 76-80.

Rigby, S. & Rogers, M. K. The general digital forensics model. Annual ADFSL Conference on Digital Forensics, Security and Law, 2007 Arlington, Virgina.

Rigger, E. 2019. *Task Definition for Design Automation.* PhD Doctorate, ETH Zurich.

Robson, C. 2002. *Real World Research: A Resource for Social Scientists and Resources*, 1st ed., Oxford: Blackwell Publishing.

Rogers, M. K. & Seigfried, K. 2004. The future of computer forensics: a needs analysis survey. *Computers & Security,* 23(1)**,** pp 12-16.

Rughani, P. H. 2017. Artificial intelligence based digital forensics framework. *International Journal of Advanced Research in Computer Science,* 8(8)**,** pp 10-14.

Sachowski, J. 2019. *Implementing digital forensic readiness: From reactive to proactive process*, 2nd ed., Boca Rato, London, New York: CRC Press.

Satchell, P. 2018. *Innovation and automation*, 1st ed., London: Routledge.

Saunders, M., Lewis, P. & Thornhill, A. 2019. *Research methods for business students*, 8th ed., United Kingdom: Pearson Education Limited.

Schober, P., Boer, C. & Schwarte, L. A. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia,* 126(5)**,** pp 1763-1768.

Seppelt, B. D. & Lee, J. D. 2019. Keeping the driver in the loop: Dynamic feedback to support appropriate use of imperfect vehicle control automation. *International Journal of Human-Computer Studies,* 125**,** pp 66-80.

Showkat, N. & Parveen, H. 2017. Non-probability and probability sampling. *Media and Communications Study***,** pp 1-9.

Shrivastava, G., Sharma, K., Khari, M. & Zohora, S. E. 2018. Role of cyber security and cyber forensics in India. *Handbook of Research on Network Forensics and Analysis Techniques.* Hershey, PA, USA: IGI Global.

Simou, S., Kalloniatis, C., Gritzalis, S. & Mouratidis, H. 2016. A survey on cloud forensics challenges and solutions. *Security and Communication Networks,* 9(18)**,** pp 6285-6314.

Suzuki, Y. 2017. Multilayered center-of-pressure sensors for robot fingertips and adaptive feedback control. *IEEE Robotics and Automation Letters,* 2(4)**,** pp 2180-2187.

Tang, T. C. 2019. *Universal Programmable IR Remote controller for home automation.* Bachelor's Electrical Engineering Degree, City University of Hong Kong.

Taylor, R. W., Fritsch, E. J., Liederbach, J., Saylor, M. R. & Tafoya, W. L. 2019. *Cyber crime and cyber terrorism*, 1st ed.: Pearson New York, NY.

Tonye, W. S. 2018. Cyber Forensic and Data Collection Challenges in Nigeria. *Global Journal of Computer Science and Technology,* 18(3-G)**,** pp 1-5.

Vallor, S. 2017. AI and the Automation of Wisdom. *In:* Powers, T. M. (ed.) *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics.* Cham: Springer International Publishing.

Verma, R., Gupta, G. & Chang, D. 2018. *Digital Forensics 2.0: An automated, efficient, and privacy preserving digital forensic investigation framework.* PhD thesis, IIIT-Delhi.

Wantoo, A., Bansal, R. & Kushwaha, P. 2017. A Review on Robotic Process Automation. *Journal of Advancements in Robotics***,** pp 23-26.

Wylot, M., Hauswirth, M., Cudré-Mauroux, P. & Sakr, S. 2018. RDF data storage and query processing schemes: A survey. *ACM Computing Surveys (CSUR),* 51(4)**,** pp 1-36.

Yadav, S. 2020. Cyber Forensics: Its Importance, Cyber Forensics Techniques, and Tools. *Critical Concepts, Standards, and Techniques in Cyber Forensics.* IGI Global.

Yamane, T. 1967. *Statistics, an introductory analysis*, 1st ed.: Harper & Row.

Yaqoob, I., Hashem, I. A. T., Ahmed, A., Kazmi, S. M. A. & Hong, C. S. 2019. Internet of things forensics: Recent advances, taxonomy, requirements, and open challenges. *Future Generation Computer Systems,* 92**,** pp 265-275.

Yaya, J. A. 2014. *Choosing the right measurement instrument for your project: Tips to apply* [Online]. Available: https://nairaproject.com/blog/measurement-instrument.html [Accessed 10 June 2022].

APPENDIX 1 – SURVEY INVITE

Department of Information Systems
Leslie Commerce Building
Engineering Mall, Upper Campus
OR
Private Bag. Rondebosch 7701
Tel: +27 (0) 21 650 4028 Fax: +27 (0) 21650 2280
Internet:
http://www.commerce.uct.ac.za/informationsystems/

Dear Sir/Madam,

You, as part of an expert sample, are invited to participate in a research study pertaining to Understanding Factors Affecting Automation in Cyber Forensics. The research is being conducted by Dean Hayes in fulfilment of a master's degree in information systems at the University of Cape Town. Please find the link to the survey at the bottom of the page.

Background

Automation has played a huge role in society by making processes easier and more efficient and has indicated a large amount of effort invested in automating the full cyber forensic process. One of the difficulties surrounding automation of this process pertains to the nature and speed of technological development. The forensic process is therefore constantly under review.

The researcher's intention is to identify the factors that will influence automation in the cyber forensics phases as it will improve industry knowledge and ability to identify barriers and inefficiencies to the process. There is a need to understand the process and in turn, it needs to be efficient enough to keep up with technological advances, reduce costs and avoid unnecessarily labour-intensive procedures. Ultimately, contributions will improve the quality of the evidence for it to be admissible in a court of law.

Your participation in this research is voluntary. All information will be treated in a confidential manner and used exclusively for the purpose of this study. No individual names will be recorded or published. You will not be requested to supply any identifiable information, ensuring anonymity of your responses. You can choose to withdraw from the research at any time for whatever reason, in accordance with ethical research requirements. This research study has been approved by the Commerce Faculty Ethics in Research Committee.

The findings of the research will be presented in a report to the University of Cape Town. The findings may also be published in an academic journal or in a conference paper if deemed to be of academic value. A copy of the report may be made available for all participants to examine. This survey uses a snowball sampling method. Please forward the survey to any other professional who you believe could add value.

Should you have any questions regarding this research, please feel free to contact me on +27 82 514 5401 or email: hysdea002@myuct.ac.za

Your participation in this study would be greatly appreciated but is entirely voluntary.

Kind regards,

Dean Hayes

Masters Student
Department of Information Systems
University of Cape Town

Email: hysdea002@myuct.ac.za

Prof. Michael Kyobe

Research Supervisor
Department of Information Systems
University of Cape Town

Email: michael.kyobe@uct.ac.za

Follow this link to the Survey:

Take The Survey

Or copy and paste the URL below into your internet browser:
https://ucpcommerce.eu.qualtrics.com/jfe/form/SV_4UvjjGIc6AhS6i1?Q_DL=eLHw0qpoYDYK6lT_4UvjjGIc6AhS6i1_MLRP_9n0VioYHPTF1UNL&Q_CHL=email

APPENDIX 2 – QUESTIONNAIRE

SURVEY : AUTOMATION IN CYBER FORENSICS

This questionnaire is confidential therefore, no one will know what you have answered. It is anonymous, so please do not put your name on it anywhere. Please note that this questionnaire is completely voluntary, and you can decide to exit at any time.

SECTION A: DEMOGRAPHIC INFORMATION

1. What is your current age group?

- Younger than 21 years (1)
- 21 to 30 years (2)
- 31 to 40 years (3)
- 41 to 50 years (4)
- Older than 50 years (5)

2. What is your highest educational level completed?

- Lower than Gr.12 (1)
- High School Certificate/ Matric (2)
- Vocational/ Higher Certificate (3)
- Diploma (4)
- Undergraduate Degree/BTech (5)
- Postgraduate degree (6)

4. What is your current employment status?

- Full-time (1)
- Part-time (2)
- Self-employed (3)
- Unemployed (4)
- Retired (5)

3. In which province(s) are you currently employed? *(Multiple selection possible)*

- Eastern Cape (1)
- Free State (2)
- Gauteng (3)
- KwaZulu-Natal (4)
- Limpopo (5)
- Mpumalanga (6)
- North-West (7)
- Northern Cape (8)
- Western Cape (9)
- Not applicable (unemployed/retired) (0)

5. Which best describes your current title?

- Forensics Investigator / Professional (1)
- Forensics/ Security Manager (2)
- Forensics Lawyer (3)
- Security Specialist/ Professional (4)
- Not applicable (0)

6. How many years of experience do you have in this role?

- Less than one year (1)
- 1 to 2 years (2)
- 3 to 5 years (3)
- 6 to 9 years (4)
- 10 years or more (5)
- Not applicable (0)

7. Which industry do you currently work in?

- Accounting and legal (1)
- Business services (2)
- Construction, repair and maintenance (3)
- Finance (4)
- Health Care (5)
- Information Technology (6)
- Manufacturing (7)
- Media and Advertising (8)
- Restaurants, bars and food services (9)
- Retail (10)
- Other (99)
- Not applicable (0)

SECTION B: AUTOMATION OF CYBER FORENSICS

1. Do you use automation during your cyber forensics investigation process?

- Yes (1)
- No (0)

2. Do you conduct automation of cyber forensics investigation in the follows phases? **SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4], SA = Strongly Agree [5]**	**SD**	**D**	**N**	**A**	**SD**
Preparation & Identification Phase	SD	D	N	A	SD
Collecting & Preservation Phase	SD	D	N	A	SD
Examination & Analysis Phase	SD	D	N	A	SD
Presentation Phase	SD	D	N	A	SD

3. When applying automation to the forensics investigation process, is the process faster/more efficient?

- Yes (1)
- No (0)

SECTION C: FACTORS INFLUENCING AUTOMATION OF CYBER FORENSICS INVESTIGATION

Independent variables

1. Indicate whether you SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4], SA = Strongly Agree [5] with the following statements based on your experience within the cyber forensics investigation.	**SD**	**D**	**N**	**A**	**SD**
a. My organisation has measures to enhance accessibility of data which helps to improve automation of cyber forensics investigation during the preparation and identification phase. [Accessibility to data]	SD	D	N	A	SD
b. My organisation has tools and systems in place to deal with challenges associated with unstable data in order to help the automation of cyber forensics investigation during the preparation and identification phase. [Unstable data]	SD	D	N	A	SD
c. My organisation uses Trust Cloud to facilitate the automation of cyber forensics investigation during the collection and preservation phase. [Dependency on Cloud Service Providers (CSPs)]	SD	D	N	A	SD
d. My organisation uses Digital Forensics as a Service (DFaaS) to facilitate the automation of cyber forensics investigation during the collection and preservation phase. [Minimize time, maximize coverage]	SD	D	N	A	SD
e. My organisation uses Push Button Forensics (PBF) such as EnCase, Forensic Tool Kit, BelkaSoft and Autospy Forensic Browser to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Decline in expert knowledge]	SD	D	N	A	SD
f. My organisation has sufficient hardware requirements such as memory size, central processing power and disk space to enable the automation of cyber forensics investigation during the examination and analysis phase. [Hardware requirements]	SD	D	N	A	SD
g. My organisation has the competencies to conduct profiling and event construction to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Profiling and event construction]	SD	D	N	A	SD
h. My organisation has put Artificial Intelligence (AI) systems to facilitate the automation of cyber forensics investigation during the examination and analysis phase. [Artificial Intelligence]	SD	D	N	A	SD
i. My organisation has expert investigators that help to facilitate the automation of cyber forensics investigation during the presentation phase. [Non-expert investigator]	SD	D	N	A	SD
j. My organisation has tools and systems that help to collect data that has been spread over the cloud to facilitate the automation of cyber forensics investigation during the presentation phase. [Spread of data in the cloud]	SD	D	N	A	SD
k. My organisation has put measures to overcome the reliability and privacy concerns that impedes the automation of cyber forensics investigation during the presentation phase. [Reliability and privacy]	SD	D	N	A	SD

2. Which of the following factors to you deem the most important to your organisation.

- Accessibility to data (1)
- Unstable data (2)
- Dependency on Cloud Service Providers (CSP) (3)
- Minimize Time, maximize coverage (4)
- Decline in expert knowledge (5)
- Hardware requirements (6)
- Profiling and event construction (7)
- Artificial Intelligence (8)
- Non-expert investigator (9)
- Spread of data in the cloud (10)
- Reliability and privacy (11)

SECTION D: AUTOMATION CYBER FORENSICS INVESTIGATION PERFORMANCE

Dependent variable

1. Indicate by whether you SD= Strongly Disagree [1], D = Disagree [2], N = Neutral [3], A = Agree [4] or SA = Strongly Agree [5] with the following statements.	**SD**	**D**	**N**	**A**	**SD**
a. Since automating cyber forensics investigation there has been a decline of cyber security incidents.	SD	D	N	A	SD
b. There has been an increase in the number of successful prosecutions as a result of automation of cyber forensics investigation.	SD	D	N	A	SD
c. Since automating cyber forensics investigation there has been an increase of productivity time and a decline in IT operational costs	SD	D	N	A	SD