International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. , No. (Mon-20..)

http://dx.doi.org/10.12785/ijcds/XXXXXX

# Building A Resilient Data Protection Awareness Framework For Data Privacy:
## A South African Case Study

*Received Mon. 20, Revised Mon. 20, Accepted Mon. 20, Published Mon. 20*

**Abstract:**
Due to the proliferation of information technology, individuals are now able to engage globally. This has exponentially increased the risk of personal information being exposed. Consequently, organizations need to implement data protection awareness to safeguard their information. Data protection is vital in all industries, and improving awareness training can potentially enhance the security of sensitive information against loss, corruption, or damage. The aim of this study is to develop a comprehensive framework for data protection awareness, with the goal of enhancing data security and privacy within organisations. The level of data protection awareness among employees of a public organisation was assessed using a quantitative method and survey research strategy. Data were collected from participants before and after they were shown an explanatory poster to gauge their current understanding of data protection. After viewing the poster and participating in a video discussion, participants were given a second questionnaire to assess their improved understanding of data protection. By comparing the results of the two surveys, the study identified areas where participants had misconceptions about data protection. The study recommends organisations to continually promote data protection awareness, especially in hybrid working environments, where an understanding of the eight conditions of POPIA is crucial.
**Keywords:** Cybercrime, Data privacy, Data protection, Employee awareness, Compliance, Hybrid working, Information security, POPIA

## 1. INTRODUCTION

This study examines the efficacy of data protection awareness in a public organisation. Employees are crucial in promoting data protection awareness in any organization. Therefore, the main goal of this research is to evaluate how effectively employees apply their data protection awareness in a hybrid work environment. The implementation of the Protection of Personal Information Act (POPIA) coincided with the South African government's initial lockdown in response to the global pandemic. This situation introduced various complexities, including the need to adapt to the new practice of remote work. The organisation committed substantial time, financial, and human resources to propagate data protection awareness in the course of the POPIA rollout. Despite the extensive efforts and investments made, it remains essential for employees to fully understand their responsibilities in implementing the data protection training they have received when dealing with personal information. Employees who directly handle data collected by the organisation are instrumental in ensuring compliance with the Act and effectively implementing data protection measures in all assigned tasks. The ability to access information and knowledge significantly affects an employee's capacity to execute data protection awareness effectively.

To raise awareness, Sayers [1] suggests a seven-step process for transferring information, starting with identifying the subject, creating employee interest, building skills and optimism (belief in success), providing a facilitation process, incorporating simulations to encourage action, and reinforcing the initial message. Despite the belief held by many that using private browsing or blocking cookies can protect them from involuntary data collection, the threat of privacy infringement persists [2]. Two significant threats to data privacy have emerged. The first is the growth of the internet and social media, which increase the power of data protection/surveillance [3], leading people to disclose private information knowingly or unknowingly. The second, more perilous threat, is the increasing importance and reliance on information when making decisions. Mason [4] asserts that policymakers consider information to be of great value, even if it means compromising the privacy of others. Information is frequently utilised, stored, shared, and reused among numerous employees or business partners in many organisations. As a result, data collection has transformed numerous organisations, transitioning from traditional verbal and written contracts to the establishment of Information Systems designed to gather data and facilitate contractual transactions.

The largest technology firms, including Apple, Microsoft, Amazon, Google (Alphabet), and Facebook (Meta), have transformed data into an asset, despite the Generalised Agreement on Accounting Principles (GAAP), which does not recognise digital personal data as an asset on balance sheets [5]. Understanding the importance of data, its protection, and its value in the wrong hands is crucial. As collectors and custodians of personal data, organisations have a responsibility to ensure secure management of data [4]. The value of data is not limited to organisations providing products and services to consumers, as it has shown to be of significant value to cybercriminals [6], [7], [8]. Irresponsible processing and management of data can have detrimental effects on both the data owner and processor [9]. This study aims to evaluate the application of data protection by an organisation's employees in a dynamic working environment. The research methodology employed a positivist philosophical paradigm and a quantitative survey research strategy.

### A. Problem Statement

The data protection awareness extends beyond South African organisations, as data security has emerged as a worldwide concern influenced by diverse political and social factors. With the increasing amount of data being shared amongst employees within an organisation, it has become crucial to raise awareness about the importance of protecting it. The primary objective of this study is to assess the effectiveness of data protection awareness training delivered to employees in a hybrid work setting, with specific emphasis on the implementation of the POPIA project.

### B. Research Objective

The overarching objective of this investigation is to attain an all-encompassing understanding of employees' cognition regarding their responsibility and accountability for compliance in a hybrid work environment, subsequent to having undergone a one-time training session during the implementation of the POPIA project. To realize this goal, a quantitative survey methodology has been employed to collect valuable feedback. Furthermore, to augment data protection awareness across all working environments, a poster (Figure 1) and a video featuring data protection insights were furnished to employees. The secondary objectives outlined below serve to enhance and reinforce the primary objective of this research study.

The objective of the pre-questionnaire is:

- to examine the existing level of data protection awareness among employees following their completion of training.

The objective of the post-questionnaire are:

- Assess the employee's comprehension and interpretation of the POPIA conditions.

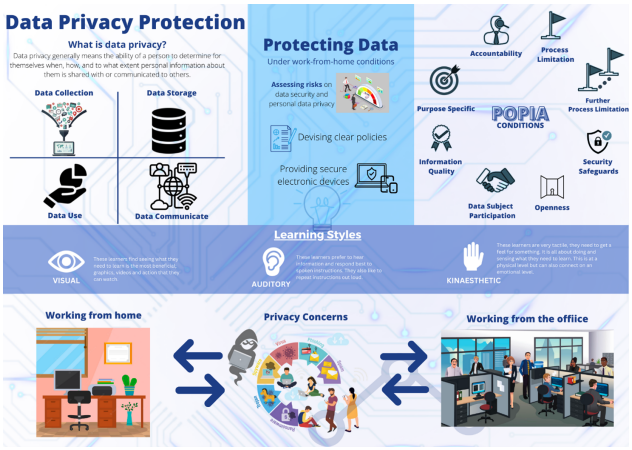


Figure 1. The data protection poster

- Evaluate the employee's level of awareness of the organisational policies and procedures concerning data protection.

- Investigate how employees apply their data protection knowledge in a hybrid working environment under various processing conditions.

The diagram presented in Figure 2 showcases how employees are implementing data protection awareness.

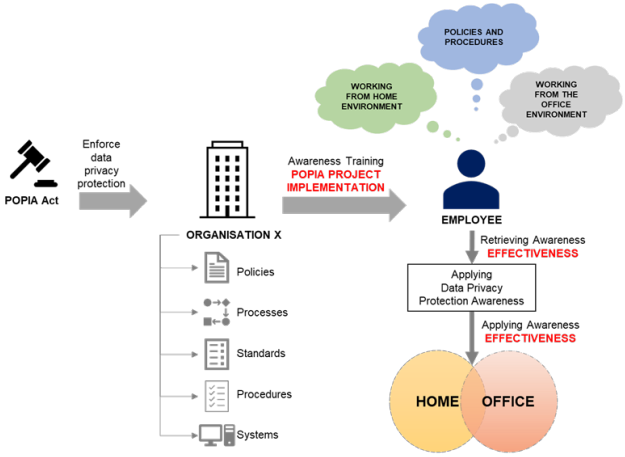


Figure 2. Assessment of the effective application of data protection awareness

### C. Research Questions

In order to accomplish the goals outlined for this study, the following primary research question has been formulated:

- In what manner do employees utilise their received data protection awareness training to adhere to POPIA regulations when processing personal information?

The secondary research questions were derived from the main research question as follows:

- Pre-questionnaire:
  - How prevalent is the current awareness of data protection?
- Post-questionnaire:
  - Did the implemented demonstration effectively enhance data protection awareness?
  - What is the extent of employees' comprehension regarding the 8 conditions outlined in the POPIA?
  - Have the organisational policies and security procedures aided employees in implementing data protection measures effectively?
  - How does the work environment of employees contribute to ensuring compliance with the POPIA through the application of their data protection awareness?

The aim of this research, as shown in Figure 3, is to explore how employees apply their data protection awareness effectively in both remote and on-site working environments. Furthermore, the study seeks to incorporate the employee's understanding of data protection after undergoing training.

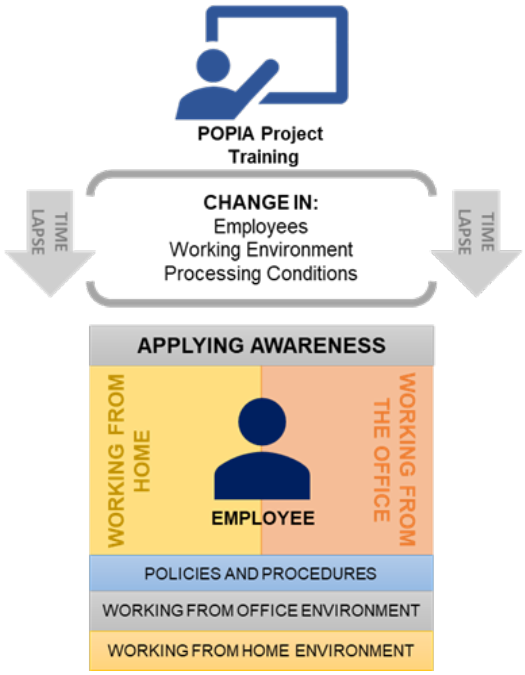


Figure 3. Research scope overview

### D. Research Scope and Targeted Community

The targeted population for this study comprises employees who regulate and supervise the financial sector in South Africa, encompassing all financial institutions providing financial products or services as defined in the Financial Sector Regulation (FSR) Act. The organization has previously conducted a successful POPIA project and provided a one-time data protection awareness and readiness training. However, changes in the employees' working environment and processing conditions have occurred since the training was conducted. Employees handle personal information regularly, and due to the COVID-19 pandemic and lockdown, a hybrid work environment has emerged where some employees work from home permanently, some work from the office, and some work from both environments periodically. The physical office environment is structured to support employees in applying their data protection knowledge. The organisation's office premises have access control measures in place to restrict data access to unauthorised employees. Personal data is handled with caution during tasks such as printing, viewing, and destruction. The organisation has also implemented additional measures to protect data, including policies, firewalls, and promoting cybersecurity awareness. The target community for this study is homogeneous and shares the following similar attributes:

- Not all participants are members of the senior management team.
- They share, collect, retrieve and/or record information.
- All participated in the awareness training.
- All work remotely.

The document is structured as follows: The literature review is discussed in Section 2, followed by an introduction to the research methodology in Section 3. The research findings are presented in Section 4, followed by a discussion and critical evaluations of the findings in Section 5. The study's limitations are addressed in Section 6. Finally, the article concludes in Section 7.

## 2. LITERATURE REVIEW

The section presents a literature review that pertains to the efficacious promotion of data protection awareness. This section is delineated into two subsections. The first subsection elucidates the concept of privacy, data privacy, and information privacy. The second subsection centres on South African data protection guidelines, which encompasses data protection legislation and data security apprehensions.

### A. The Privacy Concept

Concepts such as confidentiality, secrecy, concealment, and protection may lead to confusion when it comes to privacy [10]. According to the Cambridge English Dictionary, privacy refers to an individual's right to keep their personal affairs and relationships confidential [11]. It is a desired state of personal space, information, and thoughts that are not intended for sharing.

While there is no legal definition of data privacy, it is crucial in preventing cybercriminals from using data maliciously [12], [13], [14]. Despite the lack of a clear definition of privacy, there is an increasing agreement that privacy encompasses various facets. Scholars in privacy have sought to develop classifications of privacy concerns and categories of intrusion. As an example, Finn and Wright [15] proposed a taxonomy of seven types of privacy, which are outlined in Table I.

In his work, Alan Westin [16] provides a comprehensive analysis of the conflict between privacy and surveillance in contemporary society. He discusses the role of privacy in society and outlines its functions. Although there is a distinction between the definitions of data and information, these terms are often used interchangeably. The field of data protection can be categorised into three main areas: traditional data protection, data security, and data privacy (Figure 4).

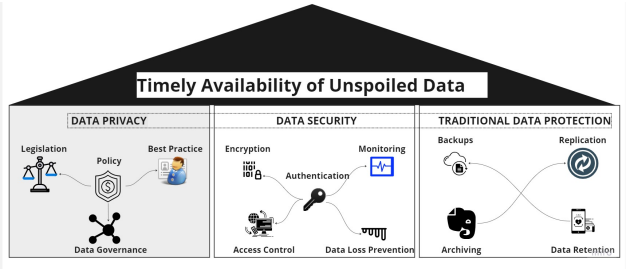


Figure 4. Data privacy, security, and protection

Data refers to discrete pieces of information that may or may not be directly associated with a particular individual [17]. Data can possess both quantitative and qualitative characteristics [18]. However, information is the result of combining various data attributes, which provides further knowledge and insight into a particular subject [17], [18]. The definitions of data and information privacy differ, therefore, the challenges around privacy are diverse and numerous. Nonetheless, the management of personal information is at the core of these definitions, with only small differences, according to Belanger et al. [19] and Rossi [18].

### *B. Data Protection Guidelines*

This section explores various South African data protection guidelines as outlined in the legislation. However, it is important to recognise that data protection awareness and application should not be limited to legal requirements. As Doctorow [20] points out, advances in technology and the internet continue to create opportunities for exploitation of information and data, despite regulatory oversight. The amount of personal data stored by individuals and organisations is increasing rapidly on a global scale, leading to greater risks of data breaches and other unlawful activities. Privacy concerns were raised by Jordaan [21] about the collection of personal information and the management of biases, unauthorised access, and secondary use of sensitive data. While the enforcement of information protection legislation is in place, some government bodies are still exempt from data protection [22]. This exemption raises concerns about finding a balanced position between the human right to privacy and the protection which governments offer to their citizens [23]. In addition, South Africa has seen a surge in cyber-attacks, with almost 60% of the population using various devices online [24], [25]. Cybercriminals primarily target critical systems to gain access to data and relevant information from organisations. Therefore, it is essential to implement effective data protection measures to safeguard against such risks.

The information in question pertains to the customers, stakeholders, partners, or employees of the organisation. Hackers are often driven by financial gain or the desire to cause harm. In South Africa, the cost of cybercrime victimisation for the public is nearly R2.2 billion per year, making the country the second-highest for cybercrime in Africa [26]. The harm and risks associated with cyber-attacks extend beyond financial loss and can have political or military motivations [27]. Due to existing cyber security threats in South Africa, there is an increasing need for data protection awareness. As per the top 10 index of countries globally afflicted by cybercriminal activities, South Africa has been listed in the fifth position (Figure 5) [28]. The index is derived from the ratio of cybercrime victims per 1 million internet users, as reported to the Federal Bureau of Investigation (FBI) from 2021 to 2022 [1].

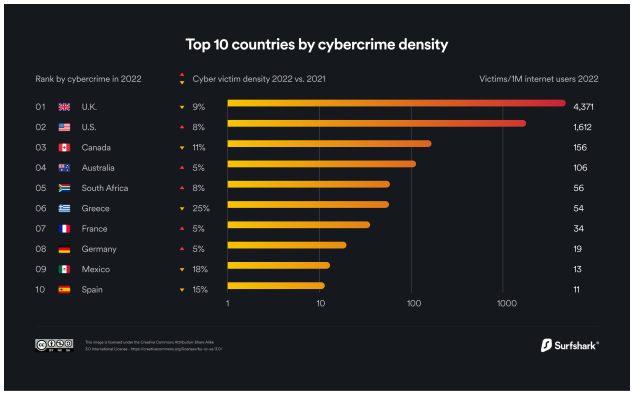


Figure 5. Top ten countries by cybercrime density

The recent cyber-attacks in South Africa include the Dis-Chem and TransUnion breaches. Consequently, data protection awareness assumes significant importance in safeguarding the confidential information of stakeholders, partners, employees, and customers from the perils of cyber-attacks. The Dis-Chem attack affected over 3.6 million South African records, while 54 million personal records were compromised in the TransUnion attack, both orchestrated by the hacker group N4aughtysecTU [2].

[1] https://surfshark.com/research/data-breach-impact/statistics
[2] https://www.itweb.co.za/content/PmxVE7KEABOqQY85

TABLE I. Privacy types (Finn and Wright [15])

| Number | Type of Privacy |
|---|---|
| 1 | A privacy related to a person<br>A privacy related to an entity |
| 2 | Privacy related to behaviours and actions |
| 3 | Privacy related to communication |
| 4 | Privacy related to data and images |
| 5 | Privacy related to thoughts and feelings |
| 6 | Privacy related to space and location |
| 7 | Privacy related to associations<br>Privacy of groups |

Furthermore, several other South African organisations have suffered from cybercrimes, such as Experian, City of Johannesburg, Ster-Kinekor, FNB, Standard Bank, ABSA, Telkom, the Department of Justice, Transnet, Uber Hack, and Momentum Metropolitan [29]. These incidents indicate a pressing need for increased data protection and security measures in South Africa. A 2017 study based on South African data showed that the most common types of cyber-attacks are data exposure and financial theft [29]. Van Niekerk [24] analysed 54 cyber-incidents and highlighted the potential impacts of Denial of Services (DoS), defacement, data corruption, system penetration, financial theft, and data exposure. The study found that data exposure and data corruption contributed to 45% of the impact of cyber incidents (Figure 6). Moreover, over half of cyber-attacks incidents in South Africa, 54% targeted state or political entities [24]. Organisations typically prioritise different aspects of data protection [30]. According to Netshakhuma and Nkholedzeni [31], public entities in South Africa generate information and data that are necessary for managing compliance with the country's laws and regulations.

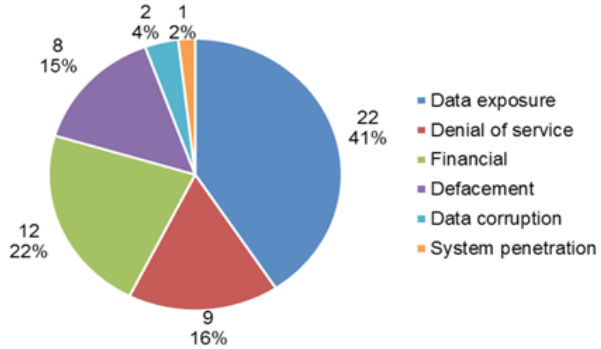


Figure 6. Analysis of cyber incidents in South Africa (Niekerk [24])

In this research, the organisation under investigation falls within the same scope outlined by Netshakhuma and Nkholedzeni [31]. Data protection is critical to the organisation, and it is essential to comply with the POPIA regulations since it collects and processes customer information. The organisation under study has implemented various measures to ensure that data protection is in line with POPIA requirements, and additional measures have been put in place to secure personal information. Figure 7 illustrates a comparison between POPIA compliance and the factors influencing the new normal for South Africans working and learning in a hybrid environment.

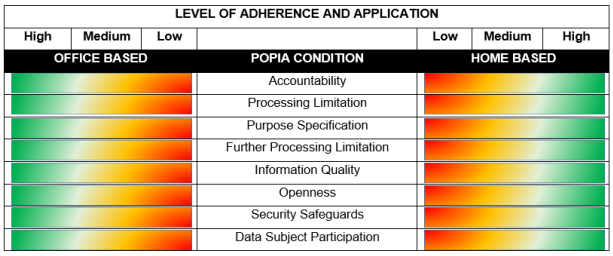


Figure 7. Employee's adherence and application of personal information

The comparison includes:

- Employee's lack of understanding about complying with POPIA.
- Ensuring compliance regardless of the environment.
- Compliance with organisational POPIA policies, standards, and procedures during data protection application.

The result presented in this study evaluates whether the measures implemented during the organisation's POPIA project have effectively prepared its employees to handle personal data protection requirements and concerns in accordance with the POPIA.

### *C. The South African Data Privacy Legislation*

The POPIA was introduced by the South African Department of Justice and Constitutional Development on 14 December 2018, with a grace period of one year for organisations to comply with its requirements from the date it became effective [32]. In preparation for its enforcement, many organisations established policies and codes of conduct as guidelines and standards to emphasise the significance of information protection.

After extensive discussions and consultations, the POPIA was promulgated by the president and became effective on 1 July 2020 to ensure that data protection is ethically managed [32]. During the Coronavirus pandemic, South Africa declared a National State of Disaster period from 15 March 2020 to 4 April 2022, which led to a nationwide lockdown with different alert levels. This period coincided with many organisations in South Africa implementing the POPIA to ensure compliance while also adapting to new work arrangements such as office-based, home-based, or hybrid models. The implementation of these new work arrangements played a critical role in promoting data protection awareness and effectively enforcing it across organisations [33].

The main objective of the POPIA is to safeguard the security of sensitive data processed by South African organisations and regulate the flow of personal information across the country's borders. The Act lays out eight conditions as depicted in Table II [32], which necessitate data protection awareness and compliance at all organisational levels. To protect personal information, various countries worldwide have instituted measures such as the POPIA and the General Data Protection Regulation (GDPR), as highlighted by David Banisar [34]. This development stems from the fact that data protection has become an essential aspect of human rights and privacy. Figure 8 illustrates the number of countries that have fully enacted data protection, those with pending enactment, and those with no data protection initiatives as of 2023, such as the Democratic Republic of the Congo.

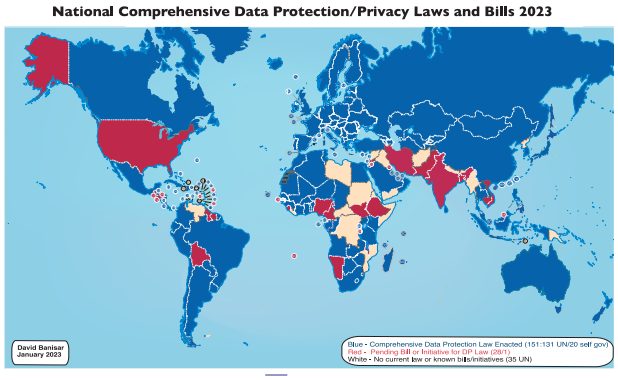


Figure 8. National comprehensive data protection/privacy laws and bills (Banisar [34])

In order to safeguard individuals' privacy, numerous data protection laws have been established globally. Banisar [34] highlights the international advancements made between December 2016 and August 2023. As demonstrated in Figure 8, it is apparent that the global data protection framework is changing. In addition to the POPIA, South Africa has enacted several laws and regulations to safeguard privacy, including:

- The Constitution of the Republic of South Africa guarantees the right to privacy, and Chapter 2 of the Bill of Rights further affirms this right.
- Section 14 of the Bill of Rights specifies that every South African has the right to privacy and sub-section (d) states the right to not have their communications privacy infringed upon [35].
- The South African Electronic Communications and Transactions Act (ECTA) of 2002 contains provisions that regulate the electronic management of communication, including data protection principles that data collectors must adhere to as outlined in Chapter 8 of the constitution [35]. Chapter 9 of the ECTA specifically addresses the protection of critical data related to the republic and the well-being of its citizens.
- The Cybercrimes Act No. 19 of 2020 was enacted to address cybercrime offences and impose corresponding penalties [35]. Chapter 13 of the Electronic Communications and Transactions Act (ECTA) outlines the following criminal offenses related to cybercrime:
  - Unauthorized access to data (e.g., hacking).
  - Tampering with website data by exploiting database vulnerabilities.
  - Inserting viruses into systems.
  - Engaging in fraud, extortion, or forgery for financial gain.

The Cybercrimes Act provides guidelines for addressing cybercrimes and malicious communications of private data, as outlined in Chapter 2 of the Act. This study highlights that the organization's employees have access to information and security policies that guide them in managing, processing, and protecting data in accordance with the protection guidelines established by these Acts.

### *D. Data and Hybrid Work Settings*

The idea of working from home or freelancing has existed for many years. However, the COVID-19 pandemic played a significant role in its widespread adoption, as global lockdowns forced organisations to implement work from home policies. Prior to the pandemic, the concept of working from home did not receive enough justification for adoption [36]. Currently, some organisations have already incorporated the hybrid work model into their structures, operations, and processes. According to Righetti et al. [37], advancements in Machine Learning (ML) and Artificial Intelligence (AI) may render some job descriptions outdated as they point towards reducing human processing. According to Thorstensson [38], when it comes to working from home there are pros and cons for both the employer and employee. The terms employee productivity and efficiency are often used interchangeably. Many organisations face difficulty in adopting the hybrid work model due to the uncertainty around maintaining productivity and efficiency [39].

TABLE II. The POPIA conditions

| Number | Condition | Expectation |
|---|---|---|
| 1 | Accountability | The party responsible for ensuring the lawful processing |
| 2 | Processing Limitations | Retention and restriction of records |
| 3 | Purpose Specific | Retention and limitation of records |
| 4 | Further Processing Limitations | Align with intended purpose of data collection |
| 5 | Information Quality | The standard of information accuracy and reliability |
| 6 | Security Safeguards | Confidentiality and integrity to maintain data security |
| 7 | Openness | Document and inform the subject about the data collection process |
| 8 | Data Subject Participation | The correction of information should be transparent, and accessible |

Additionally, employees have a responsibility to apply data protection measures whether they are working from home or at the office. There are advantages and disadvantages to working in a hybrid environment. Some of the drawbacks of a hybrid work setting include the increased risk of cyber threats and the potential for data breaches. Consequently, the objective of this research is to create a framework that promotes data protection awareness to safeguard employees of a public organisation from potential data breaches that may occur in a hybrid working environment.

## 3. RESEARCH METHODOLOGY

The aim of this research is to investigate the level of data protection awareness among employees and their corresponding responsibility and accountability towards data security. The research seeks to obtain an objective and in-depth understanding of how effectively employees apply their knowledge of data protection, particularly in relation to the training provided during the POPIA implementation project. It is worth noting that the training was delivered online while the employees were working from home due to the COVID-19 pandemic. This study is focusing on the following aspects:

- The effectiveness of the training in promoting data protection awareness among employees
- The level of employees' knowledge and understanding of data protection principles and practices
- The extent to which employees apply data protection measures in their daily work activities
- The challenges faced by employees in implementing data protection practices while working from home
- The role of organisational culture in promoting a data protection mindset among employees
- The impact of the hybrid work environment on data protection practices and awareness among employees

The research aims to contribute to the development of a data protection awareness framework that can be applied to a hybrid work environment to ensure the security of employees' data. The present study adopts a research methodology that draws from the research onion framework developed by Saunders et al. [40]. The selection of research options aligns with the layers of the research onion, as depicted in Figure 9. The chosen options are described in detail in the following subsections, which pertain to the philosophical stance, approach to theory development, methodological choice, research strategy, and time horizon. By adopting this methodological approach, the study aims to ensure a systematic and rigorous investigation of the research problem.

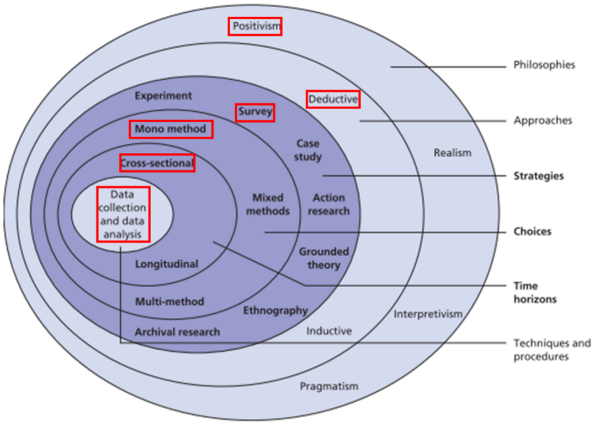


Figure 9. The research methodology (Saunders et al. [40])

### A. *Research Philosophy*

Saunders et al. [40] proposed a research onion that consists of several layers, including philosophical positions. In this context, there are four philosophical positions that researchers can adopt in their studies. The first is positivism, which is based on the belief that research should be conducted objectively, and the data collected should be empirical and measurable. The second position is interpretivism, which acknowledges the subjectivity of research and focuses on understanding the social and cultural context of the data. The third position is realism, which recognises the existence of an objective reality and aims to understand it by combining empirical data and theoretical concepts. The fourth position is pragmatism, which advocates for a combination of different philosophical positions and methods, depending on the research problem and objectives.

It is crucial to consider these philosophical positions when designing a research study, as they can influence the research approach, methods, and data interpretation [40]. In this study, a positivist research philosophy was adopted to maintain objectivity in data collection and interpretation. The researchers remained independent throughout the study by maintaining minimal interaction with the participants. The use of a positivist research philosophy was deemed appropriate as it aligns with the epistemological belief that knowledge can be obtained through empirical observation and objective [40]. This research philosophy is underpinned by the belief that reality exists independently of the individual's perception and can be studied and understood through the use of scientific methods. As such, the positivist research philosophy is well-suited to studies that seek to test hypotheses and establish causal relationships between variables through the use of empirical data.

### *B. Research Approach*

In this study, a deductive research approach based on the research onion model by Saunders et al. [40] was used. The inductive approach, as outlined by Saunders et al. [40], aims to gain insights into the meaning attached to events and provides a more flexible structure that allows for change. In contrast, the deductive approach, which was used in this research, employed questionnaires to create an understanding of observations and enable the comparison of different views of people through empirical data. The deductive approach began with a theory and moved towards hypotheses or questions, which were then tested through data collection [40]. The study followed a structured approach with controls in place to ensure the validity of the collected data. The results were analysed to confirm or reject the hypotheses.

### *C. Research Strategy*

Various research strategies exist, as outlined by Saunders et al. [40]: experimental research, survey research, case study research, grounded theory research, ethnographic research, action research, and archival research. The selection of a research strategy depends on the research question, the nature of the research problem, and the research objectives [40]. The experimental research strategy aims to test causal relationships by manipulating the independent variable and measuring the dependent variable [40]. Survey research gathers data from a representative sample to make generalisations about a population. Case study research focuses on in-depth investigation of a particular phenomenon, situation, or organisation [40]. Grounded theory research aims to develop a theory by systematically analysing data. Ethnographic research involves the immersion of the researcher in the research setting to understand the culture and behaviour of the participants [40]. Action research involves collaborating with participants to solve a specific problem. Archival research utilises data from existing records and documents to answer research questions. The selection of a research strategy is critical to ensure the research is appropriate and effective in answering the research question and achieving the research objectives.

The selected research strategy for this study is the survey approach, which involves the use of questionnaires to collect data from a sample of participants. According to Saunders et al. [40], the survey strategy is one of the various research strategies that can be used, including experiments, case studies, interviews, and systematic literature reviews. The survey strategy was chosen for this research as it provides an effective means of collecting vast and reliable data to answer the research questions [40]. To ensure the accuracy and reliability of the data collected, the participants were informed of the research purpose and the voluntary and confidential nature of the survey. Additionally, no personal identifiable information was collected from the participants, and the survey results were made available.

### *D. Research Choice*

The chosen methodological approach is a mono quantitative method, in line with Saunders research onion [40]. This approach is distinct from mixed and multi-method research, which involve a combination of quantitative and qualitative data collection and analysis techniques. The mono method can be utilized in either a qualitative or quantitative approach [40]. In this study, a mono method quantitative approach was selected to ensure objectivity and accuracy in capturing participants' responses. This approach facilitated efficient and prompt data collection.

### *E. Research Time Horizon*

As per the selected research design, the time horizon for this study is cross-sectional. The organisation had already predetermined the time frame for the data collection, and hence the chosen time horizon. According to the research onion model proposed by Saunders et al. [40], the longitudinal study time horizon is utilised to examine changes and developments over time. However, this study aimed to capture a specific momentary understanding of the employees' data protection awareness and compare it with a subsequent snapshot following the provision of additional information or training. In this regard, the cross-sectional time horizon was deemed most appropriate to obtain a one-time, instantaneous representation of the participants' perceptions and attitudes towards data protection.

### *F. Data Collection and Data Analysis*

Figure 10 depicts a visual representation of the process followed for data collection. Using Microsoft Forms, two surveys were administered to collect data for this study. The first survey aimed to evaluate participants' comprehension of data protection based on the POPIA training they had received.

The training was conducted remotely using animated videos, followed by a test focusing on a specific POPIA condition. After participants were exposed to a data protection-related poster artifact, the second survey was administered.

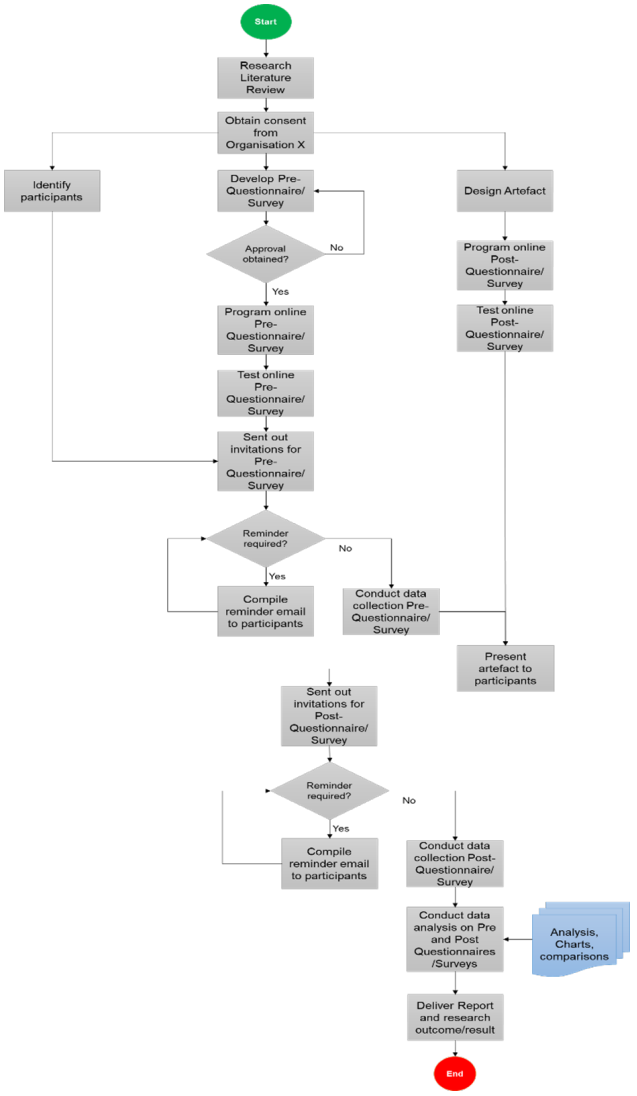


Figure 10. Outline of the research process

Additionally, a video was provided to elaborate on the content of the poster subsequent to the completion of the initial survey. The surveys were primarily composed of True or False questions, with some employing a Likert Rating scale, and others allowing for multiple answers. The questions in the second survey were primarily based on the main themes presented in the poster. The present study employed a comparative analysis to evaluate the level of awareness of data protection among participants prior to and subsequent to their exposure to a poster and its accompanying explanatory video. A summary of the research methodology implemented in this study is presented in Table III.

### G. Data Analysis Method

Ascertaining the level of data protection awareness among participants was the main objective of this study, and to this end, the authors utilised a quantitative approach to data analysis. By relying on numerical data, the authors were able to process and extract meaningful insights, which they subsequently presented using charts, graphs, and statistics. The survey questionnaire provided both opinions and factual data, the latter of which was generated through predefined answers to specific questions. An exploratory and descriptive analysis of the collected data was then conducted, enabling the authors to design and carry out research experiments, ultimately leading to a conclusion regarding the level of data protection awareness among the participants.

### H. Data Analysis and Visualization Tools

The researchers utilised Microsoft Excel and SPSS Statistics version 28 as their primary data analysis and visualisation tools. The survey responses were extracted from Microsoft Forms in Excel format, which were then imported into the SPSS software for machine-readable numeric conversion and further analysis. In order to accomplish this, coding was necessary. A codebook was developed to provide an in-depth explanation of each variable used in the study, as demonstrated in Figure 11.

**Study codebook: Pre-Questionnaire**

| Construct | Variable Name | Format | Variable Label \| (Description) \| Coded Value = Value Label |
|---|---|---|---|
| Survey Information | s1v0_id | Numeric<br>Width = 3<br>Decimals = 0 | *Survey ID*<br>(Unique identifier for the anonymous response) |
| | s1v1_consent | Numeric<br>Width = 1<br>Decimals = 0 | *Consent Given*<br>(Participants indicated their willingness to participate in study)<br><br>0 = No, I disagree<br>1 = Yes, I agree |
| Biographical Information | s1v2_bornGroup | Numeric<br>Width = 1<br>Decimals = 0 | *Year Born Group*<br>(The year group the participant was born in)<br><br>1 = 1925 - 1945<br>2 = 1946 - 1954<br>3 = 1955 - 1965<br>4 = 1966 - 1976<br>5 = 1977 - 1994<br>6 = 1995 - 1999<br>7 = 2000 - 2003 |
| | s1v2_ageGroup | Numeric<br>Width = 1<br>Decimals = 0 | *Age Group*<br>The calculated age group the participant belongs to<br><br>1 = 77 - 97<br>2 = 68 - 76<br>3 = 57 - 67<br>4 = 46 - 56<br>5 = 28 - 45<br>6 = 23 - 27<br>7 = 19 – 22 |

Figure 11. The codebook sample

### I. Population, Sampling and Response Rate

It is noteworthy to mention that the sample frame for this study comprised of 537 employees from human resources departments, selected based on their homogeneous characteristics. The targeted population for this research encompasses employees across all organisational divisions, excluding management. The utilisation of probability random sampling in this study is demonstrated in Figure 12. The employment of probability sampling as a method of data collection ensures that the research findings can be generalised to the population with a high degree of accuracy [40]. To ensure the absence of duplication in the final sample of 315 employees obtained through random sampling, each participant was assigned a unique numeric identifier using the Random (=RAND) formula in Microsoft Excel. The participants were invited to take part in the study through email, which included a link to the Microsoft Forms surveys. Before completing the survey, participants received an email detailing the purpose and nature of the study.

TABLE III. Research methodology summary

| | Research onion concepts | Research study selection |
|---|---|---|
| 1 | Philosophy | Positivist |
| 2 | Theory development | Deductivist |
| 3 | Methodological choice | Mono quantitative |
| 4 | Strategy | Survey |
| 5 | Time horizon | Cross-sectional |

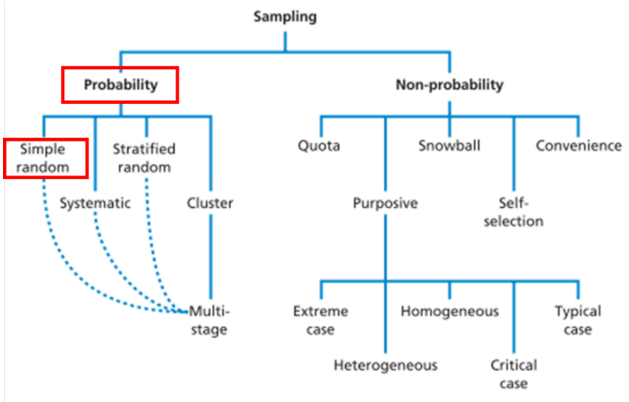


Figure 12. The sampling technique (Saunders et al.[40])

For the post-questionnaire survey, participants were provided with an attachment of the data protection awareness poster and a link to a video presentation that offered additional explanation of the poster's content. Data collection for both surveys spanned a duration of five weeks, during which multiple reminders were sent to participants to encourage their response. Various strategies were implemented to further improve the response rate.

As part of ethical considerations, the privacy and confidentiality of participants' responses were prioritised to ensure the security of the collected data. Additionally, participants were assured that the organisation had granted approval for the study, and the survey content was designed to be relatable to the participants. The surveys were crafted to be engaging and understandable, facilitating participants' understanding of the survey questions. These measures were implemented to ensure the ethical and efficient execution of the survey research. The study achieved a final response rate of 33.3% (n = 105 out of 315) for the pre-questionnaire and 31.4% (n = 99 out of 315) for the post-questionnaire, following the participants' explicit consent. As mentioned by Saunders et al. [40], non-responses are inevitable in any research study. In this study, the non-response rate was 66.6%, which included 4.44% (n = 14 out of 315) of employees who chose not to provide consent.

### J. *The Survey Design*

In the survey, there were 55 questions that were split into two different questionnaires (the pre-questionnaire and the post-questionnaire).

In addition, the survey included a demonstration tool designed to enhance data protection awareness. This demonstration was presented to the participants upon completion of the pre-questionnaire, which aimed to assess their existing knowledge and understanding of data protection. Once the demonstration tool was presented, the post-questionnaire was used to measure how well the participant understood the information. Figure 13 depicts how the survey was set up.

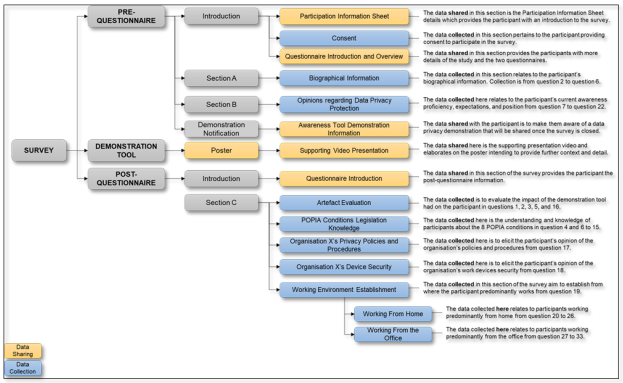


Figure 13. The survey design

## 4. RESULTS AND FINDINGS

The researchers obtained necessary approvals from the relevant authorities in the organisation, including a written approval letter and an ethical clearance certificate that were included in the permission request. The survey participants were given a detailed explanation of the research objective and purpose, along with a consent section to ensure their willingness to participate. The study results and findings will be presented in a manner that corresponds to the data protection awareness level, the participant's biographical information, and their opinions.

### A. *Findings on Biographical Information*

The initial survey encompassed a section devoted to the collection of biographical data from the study participants. Specifically, this section obtained information on the year of birth, gender, education level, first language, and employment type of the participants.

An analysis of this data reveals that the predominant proportion of participants, comprising 79.0% of the total, were born within the period between 1977 and 1994, as

evidenced by the graphical representation of this distribution displayed in Figure 14.

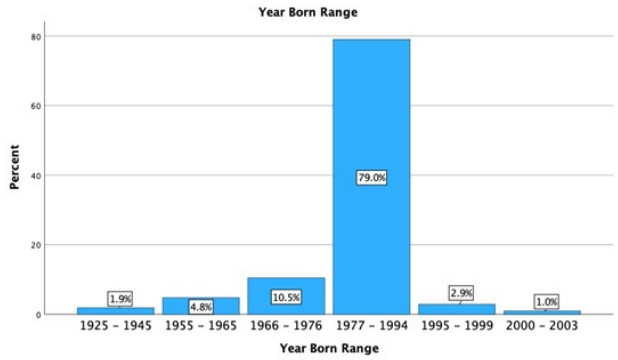


Figure 14. Distribution of the year of birth

The current study examined several demographic characteristics of the participants, such as age, gender, educational background, and primary language. The age range of the 105 participants was from 22 to 97 years old, with the majority falling between 28 and 45 years old, as indicated by 79.0% of the responses (Figure 15). Among the participants, 52.4% (n = 55) were identified as male, 44.8% (n = 47) identified as female, and 2.9% (n = 3) preferred not to disclose their gender. Regarding the highest level of education, the majority of participants held a 3-year university degree (34.2%, n = 36) or an Honors degree (32.3%, n = 34). Participants reported various education levels, including Below Grade 12 (0.9%, n = 1), Grade 12 in High School (9.5%, n = 10), Diploma (11.4%, n = 12), and Master's Degree (8.6%, n = 9). None of the participants reported having a Doctorate Degree, and 2.9% (n = 3) selected "None of the Above." It is important to note that participants were allowed to choose only one answer, and all 105 participants provided a response to this question (refer to Figure 16). Furthermore, participants were asked to indicate their first language, and the majority of participants selected Tswana (15.2%, n = 16), followed by English and Northern Sotho (13.3%, n = 14 each). The pre-questionnaire provided 11 language options for participants to choose from, including an option for "Other" (Figure 17). Additionally, the pre-questionnaire collected data on participants' employment type within the organisation. All 105 participants responded to this question, and only one answer option could be selected. Various responses were reported, including Zulu (12.4%, n = 13), Afrikaans, Venda, and Sotho at an equal proportion (8.6%, n= 9), Xhosa (7.6%, n = 8), Tsonga (6.7%, n = 7), Ndebele (4.8%, n = 5), and Other, which encompassed Siswati (1.0%, n = 1). The participants were presented with three employment options, namely, Permanently Employed, Contractually (Intern) Employed, and Temporarily Employed. Results indicated that all participants (100%, n = 105) chose Permanently Employed, as shown in Figure 18. It is noteworthy that participants could select only one answer, and all 105 of them responded to this question.

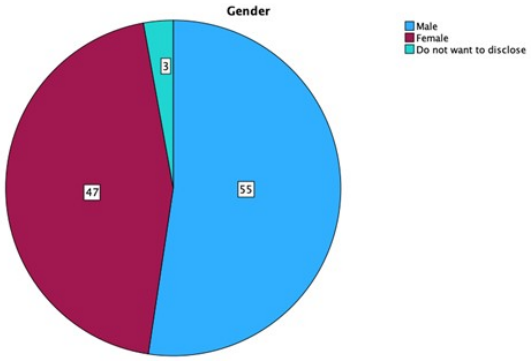


Figure 15. The gender distribution

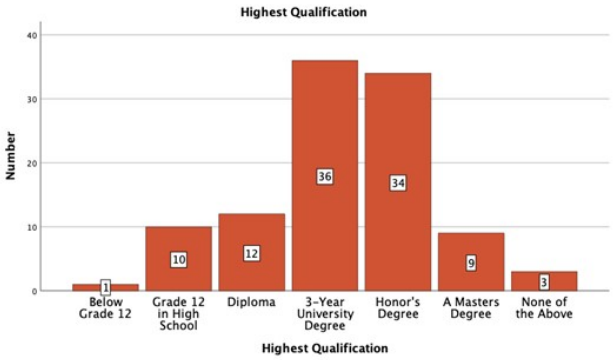


Figure 16. Distribution of the participants' highest qualification

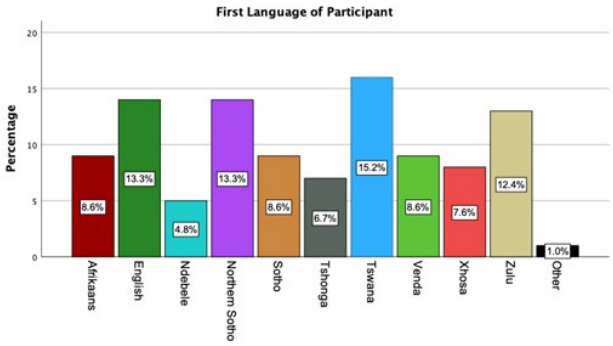


Figure 17. Distribution of the participant's language

## B. *Findings on Participants Opinions*

In the pre-questionnaire's Section B, the focus was on evaluating the participant's current awareness level, expectations, and position regarding data protection. Questions 7 to 22 were specifically designed to gauge the participant's level of exposure to data protection and to elicit their opinions on the matter.

The participant's awareness level regarding data protection was assessed using questions 7 to 11, and the responses are presented as follows:

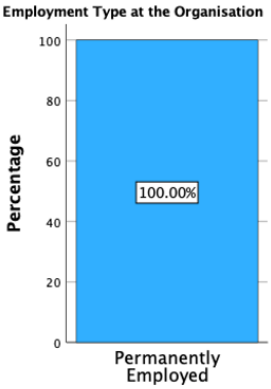


Figure 18. Distribution of the type of employment

- The primary objective of question 7 of the questionnaire was to gauge the willingness of the participant to acquire or enhance their knowledge about data protection. A total of 105 participants responded to this query, and the outcomes are presented in Figure 19.

| No. | Question | 'Yes' Response N | 'Yes' Response Percent | 'No' Response N | 'No' Response Percent |
|---|---|---|---|---|---|
| 7 | Would you want to be educated about data privacy protection awareness? | 102 | 97.1% | 3 | 2.9% |

Figure 19. Participant's willingness to be educated

- The objective of question 8 is to determine whether the participants have knowledge of the potential risks connected with data protection. The response rate was 100% (n = 105), and the findings are demonstrated in Figure 20. This inquiry was formulated to obtain information on the conversation around the potential risks linked to data protection. Similarly, the participants were limited to selecting only one answer. Out of the 62 individuals who answered affirmatively to question 8, all provided responses, and the results are presented in Figure 21.

| No. | Question | 'Yes' Response N | 'Yes' Response Percent | 'No' Response N | 'No' Response Percent |
|---|---|---|---|---|---|
| 8 | Did anyone discuss the possible dangers and threats associated with data privacy protection with you? | 62 | 59.0% | 43 | 41.0% |

Figure 20. Dangers and threats discussion with the participant

- The purpose of question 9 was to obtain details about discussions related to dangers and threats associated with data protection. This was an open-ended question where participants provided free text responses. The researchers analysed the responses and categorised them into common themes, which included the provider (who) and frequency (when) of the discussions. The themes are presented in Table IV.

| No. | If you answered "Yes", please describe in the space provided how often and who discuss these threats with you? | Response N | Response Percent | Cumulative Percentage |
|---|---|---|---|---|
| 9 | External to the organization – Unknown | 3 | 2.9% | 4.8% |
| | Internal to the organization – High | 33 | 31.4% | 53.2% |
| | Internal to the organization – Medium | 9 | 8.6% | 14.5% |
| | Internal to the organization – Low | 4 | 3.8% | 6.5% |
| | Internal to the organization – Once-Off | 3 | 2.9% | 4.8% |
| | Internal to the organization – Unknown | 10 | 9.5% | 16.1% |
| | **Total of participants that responded *Yes* to question 8** | **62** | **59.0%** | **100%** |

Figure 21. Dangers and threats discussion details

- The purpose of question 10 was to gather the participants' self-assessment of their knowledge of data protection. Each of the 105 participants provided a single response, and the resulting data is displayed in Figure 22.

| No. | Please rate your data privacy protection knowledge? | Response N | Response Percent |
|---|---|---|---|
| 10 | Very Good | 17 | 16.2% |
| | Good | 47 | 44.8% |
| | Average | 37 | 35.2% |
| | Poor | 0 | 0.0% |
| | No Awareness | 4 | 3.8% |
| | **Total Responses** | **105** | **100%** |

Figure 22. Participant's self-rating of data protection awareness

- Question 15 was designed to obtain information about the preferred medium for delivering data protection awareness material, as perceived by the participants. They were given multiple options to select from, including the Other category where they could provide their own response, such as External Media Advertising, Seminars, or Corporate Local Area Network (LAN) Advertising. All 105 participants responded to the question, and the resulting data is illustrated in Figure 23.

| No. | How would you want to be educated about data privacy protection awareness? You may select one or more options? | Response N | Response Percent | Percent of Cases |
|---|---|---|---|---|
| 15 | Brochure | 30 | 11.8% | 28.6% |
| | Workshop | 69 | 27.2% | 65.7% |
| | Online Awareness Material | 73 | 28.7% | 69.5% |
| | Poster | 20 | 7.9% | 19.0% |
| | Competitions | 21 | 8.3% | 20.0% |
| | Games | 24 | 9.4% | 22.9% |
| | Leaflets | 13 | 5.1% | 12.4% |
| | Other | 4 | 1.6% | 3.8% |
| | **Total** | **254** | **100%** | **241.9%** |

Figure 23. Participant's awareness material medium preference

- Question 19 aimed to determine whether any of the 105 participants had experienced a data breach. They were asked to choose a single response, and all participants provided an answer, as shown in Figure 24.
- Question 20 was designed to determine the type of data breach experienced by the participant.

TABLE IV. The discussion themes

| Theme | Theme Description |
|---|---|
| **Who (providers)** | |
| Internal | Conversations can take place among any member of the organisation |
| External | Conversations are conducted by an external party not affiliated with the organisation |
| **When (Frequency)** | |
| High | The discussions take place frequently, using a variety of mediums, at least five times per year |
| Medium | The discussions take place on an intermittent basis, quarterly, and through various mediums |
| Low | Various mediums are used to conduct the discussions, which occur either annually or biannually |
| Once-Off | The discussion was only held during the initiation of the organisation's POPIA project implementation |
| Unknown | The frequency and timing of the discussions were not explicitly mentioned by the participant |

| No. | Question | 'Yes' Response N | 'Yes' Response Percent | 'No' Response N | 'No' Response Percent |
|---|---|---|---|---|---|
| 19 | Have you ever been a victim of a data breach? | 28 | 26.7% | 77 | 73.3% |

Figure 24. Participant's exposure to a data breach

The responses were classified into specific themes related to common types of data breaches, as determined by this study. Table V summarises the themes identified from an open-ended question where participants were asked to describe the type of data breach they had experienced. This question allowed for free-text responses. All 28 participants who responded Yes to question 19 provided an answer, and the results are presented in Figure 25.

| No. | If yes, please provide more details (If you do not wish to divulge the nature of the breach, you can opt not to answer this part). | Response N | Response Percent | Valid Percent |
|---|---|---|---|---|
| 20 | Negligence | 1 | 1.0% | 3.6% |
| | Phishing | 3 | 2.9% | 10.7% |
| | Ransomware | 1 | 1.0% | 3.6% |
| | Spyware | 3 | 2.9% | 10.7% |
| | Theft/Loss | 2 | 1.9% | 7.1% |
| | Unauthorized Access | 10 | 9.5% | 35.7% |
| | Unknown | 8 | 7.6 | 28.6% |
| | **Total Responses** | **28** | **26.7%** | **100%** |

Figure 25. Type of data breach experienced

## C. Promoting Data Protection Awareness

During the data protection awareness demonstration, a section was specifically dedicated to emphasising the importance of catering to different learning styles when presenting information. Taylor et al. [41] identified four factors that can influence employees' learning style or orientation. The learning style or orientation of employees can be influenced by various factors, including those related to the external business environment and the work environment. Additionally, the learning potential of the job and the employees' own learning orientations are important considerations [41]. The effectiveness of an organisation's data protection awareness efforts may depend on its ability to cater to employees' learning styles and promote self-sufficiency in their implementation. In the post-questionnaire, participants were asked about the educational impact of a poster and supporting video presentation on data protection, which they had seen during an artifact demonstration. Question 1 aimed to gather participants' opinions on the effectiveness of the poster and video, with three response options: Yes (63.64%, n = 63), No (0.0%, n = 0), and To a Certain Extent (36.36%, n = 36). Only one response was permitted, and all 99 participants provided an answer, as illustrated in Figure 26. Question 2 focused on gathering participants' opinions about the informativeness and presentation of the artifact.

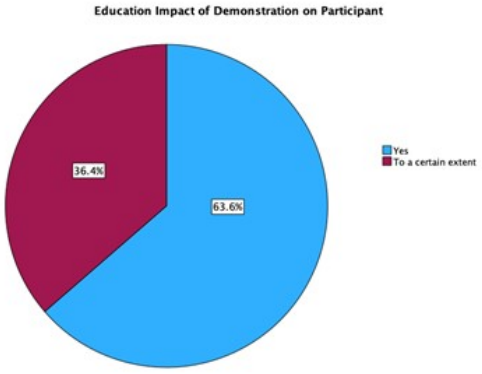


Figure 26. The educational impact of the artifact

Question 2 intended to assess the participants' opinions concerning the informativeness and presentation of the poster and supporting video. The query presented the participants with three response options: Yes (77.8%, n = 77), No (0.0%, n = 0), and To a Certain Extent (22.2%, n = 22), and only one response was allowed. All 99 participants responded, and the outcome is displayed in Figure 27. Question 3 was designed to assess the participants' basic awareness of achieving data protection. It tested if the information conveyed in the demonstration regarding achieving data protection was comprehended. Participants were given two response options: True (96.0%, n = 95) and False (4.0%, n = 4), with only one response permitted. All 99 participants responded, and the outcome is demonstrated in Figure 28. Question 5 of the survey collected data on the participant's fundamental awareness impact of data processing.

TABLE V. Data breach themes

| Theme | Theme Description |
|---|---|
| Viruses | Code that infects the computer |
| Spam | Unsolicited messages sent |
| Phishing | Fraudulent email |
| Spyware | Malware installed on the device<br>Device invalidation and data loss |
| Ransomware | Attacks denying access to files and computer<br>Ransom payment after encryption of files<br>Cybernetics attacks |
| Unauthorised Access | Illegal access |
| Negligence | A person's inability to perform data protection |
| Lack of Knowledge | Lack of data security knowledge |

The question's objective was to assess the participant's comprehension of data processing, with two options available for selection, namely True (88.9%, n = 88) and False (11.1%, n = 11).

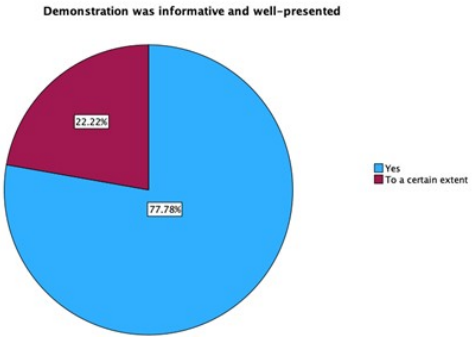


Figure 27. The informativeness and presentation of the artifact


Figure 28. Basic awareness of achieving data protection

All 99 participants responded, and the results are depicted in Figure 29. Question 16 was designed to evaluate the participant's understanding of the cardinal rule of data protection presented to them during the artifact demonstration. The objective was to determine whether the respondents had a clear understanding of the essential elements of safeguarding data. The query presented three alternatives to choose from:

Ensuring Your Environment is Secure (13.1%, n = 13), Evaluating Data Security and Data Privacy (80.8%, n = 80), and Secure Processing of Data.

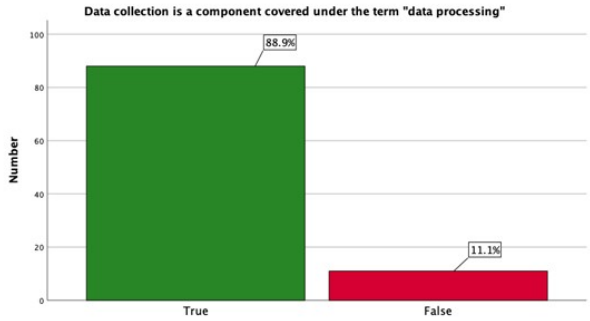


Figure 29. Basic awareness impact of data processing

While participants could only select one answer, all 99 individuals responded to the question, and the results are depicted in Figure 30. Figure 31 illustrates the assessment of the participants' comprehension and familiarity with the eight POPIA conditions categorised by their predominant work environment.


Figure 30. Fundamental rule when protecting data

### *D. Findings on Participant Opinion of the Organisation Policies and Procedures*

The focus of question 17 was to gather the participant's viewpoint regarding policies and procedures concerning

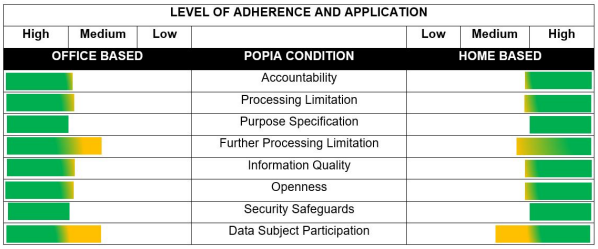


Figure 31. The POPIA awareness evaluation

data protection. The demonstration provided an overview of the role of organisational policies and procedures in data protection. The measurement of the participant's opinion was conducted through a Likert rating scale consisting of five points, ranging from strongly agree to strongly disagree, as outlined in Table VII. Although Likert scales are convenient and easy to complete, they may pose a challenge as participants may feel compelled to agree with certain statements in a manner that appears expected [42]. The response rate of question 17 was 100%, and the results are illustrated in Figure 32. Question 18 presented the findings of several statements posed to the participants to assess their opinion through a five-point Likert rating scale. Table VI outlines the statements and their corresponding rating scale (where 1 equals strongly agree; 2 equals agree; 3 neither agree nor disagree; 4 equals disagree; and 5 equals strongly disagree). The response rate for these questions was also 100%, and the results are illustrated in Figure 33, which refers to the codebook for each statement.

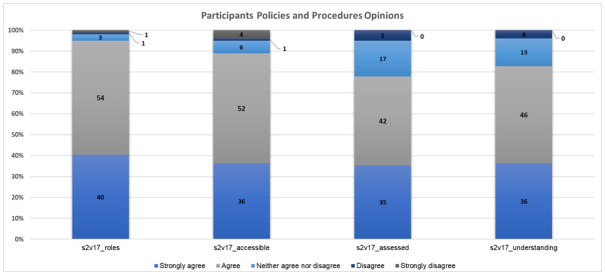


Figure 32. Participant's opinion of the organisation policies and procedures

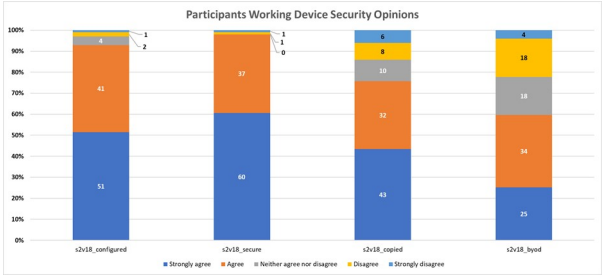


Figure 33. Participant's opinion of the organization's working device security

## E. *Findings on Participant's Working Environment*

Question 19 aimed to determine the primary working environment of the participants when processing sensitive data. The demonstration provided the participants with an overview of the environmental design in general workplaces, such as demarcated working areas based on departmental responsibilities and access controls. However, the demonstration also highlighted that such controls may be substantially decreased when working from home, which could increase the risk of data exposure. Participants were presented with two options, Working From the Office Most of the Time and Working From Home Most of the Time, and were asked to select only one answer. All 99 participants responded, and the results are shown in Figure 34. The post-questionnaire included questions (questions 20 to 26) targeted at participants who predominantly work from home. Those questions aimed at assessing the environment and gaining a comprehensive understanding of how the working environment may impact data protection.

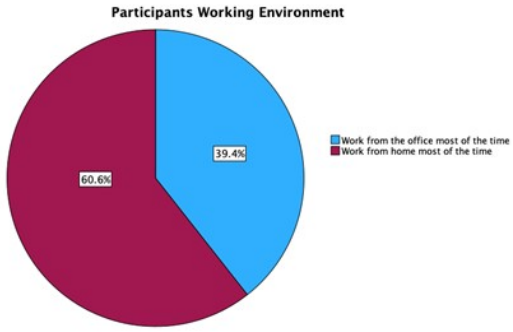


Figure 34. The participant's working environment

The survey questions 27 through 33 were specifically targeted towards participants who reported working predominantly from the office.

- Questions 20 and 27 were designed to elicit the type of data processing activities performed within a particular working environment.
  These questions were constructed as multiple-answer questions. The responses from all 99 participants are presented in Figure 35.

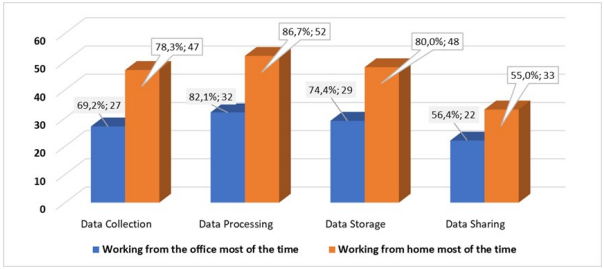


Figure 35. Type of data processing performed

TABLE VI. Codebook mapping for policies and procedures

| Codebook Variable Name | Statement |
|---|---|
| s2v18 roles | Clear roles of procedures |
| s2v18 accessibility | Data is accessible |
| s2v18 assessed | Evaluation of an employee data protection knowledge |
| s2v18 understanding | To assess an employee understanding |

TABLE VII. Codebook mapping for work device security

| Codebook Variable Name | Statement |
|---|---|
| s2v17 configuration | Accurate configuration of devices |
| s2v17 security | Devices are secure |
| s2v17 copying | Sensitive data cannot be copied |
| s2v17 byod | Bring-Your-Own-Device |

- Questions 21 and 28 intended to extract information regarding the available working tools utilised by the participant when processing sensitive data. Both of these were multiple-choice questions, and all 99 participants responded, with the results displayed in Figure 36.

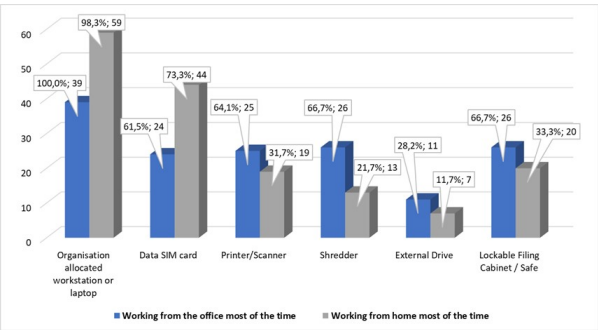


Figure 36. Type of working tools

- Questions 22 and 29 were formulated as multiple-answer questions with the purpose of identifying the type of device employed by the participant for accessing the organisation network while performing work-related activities. The responses of all 99 participants were collected and analysed. The results are depicted in Figure 37.

- Question 26 and 33 inquired from the participants whether access to sensitive data is secured. Only one option was permitted to be selected for each question, and all 99 participants responded. The outcome is depicted in Figure 38.

## 5. DISCUSSION AND CRITICAL EVALUATION

The survey achieved a response rate deemed acceptable for research studies, with 33.3% and 31.4% for the pre and post-questionnaires, respectively.

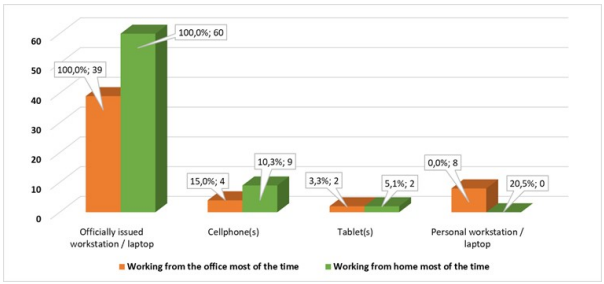


Figure 37. Type of network access device

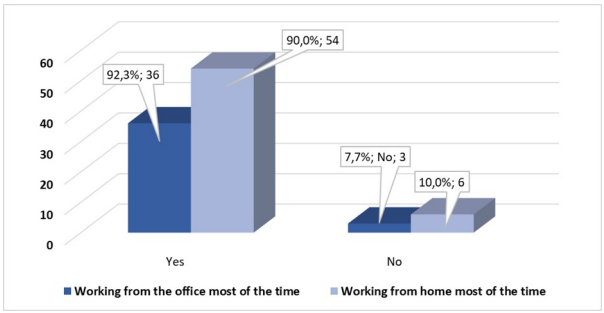


Figure 38. Physical working environment controls

The surveys were designed to only include fully completed responses and disregarded incomplete surveys, in which participants provided consent but failed to complete the survey. Based on the results of this study, it can be inferred that the sample of participants is representative of the selected population within the organisation. Furthermore, this study has identified the absence of a standardised instrument for measuring data protection awareness in a South African context. Consequently, this research has adopted an objective approach to gauge the level of data protection awareness in a South African context.

### A. Biographical Information

In the pre-questionnaire, Section A was designed to gather biographical information of the participants who were non-managerial staff of the organisation. The sampling frame exclusively comprised of non-managerial staff. The findings of the survey revealed that majority of the participants were born between 1977 and 1994 (79.0%, n = 83). Males were the most frequently represented gender (52.4%, n = 55), and the most popular qualification was a 3-year University Degree (34.2%, n = 36). Furthermore, the majority of the participants held a permanent job (n = 105) and identified Tswana (15.2%, n = 16) as their home language.

### B. Participants Opinion

One of the sub-objectives of this study was to investigate the level of data protection awareness among employees in relation to their training. To achieve this objective, a research sub-question was formulated to determine the current state of data protection awareness among employees. To answer this sub-question, data was collected through a pre-questionnaire. The findings revealed that participants' perceived level of data protection awareness proficiency was at an average confidence level. However, their willingness to be educated on data protection was high (97.1%), indicating a strong desire for more training. The responses received on discussions about data protection risks and threats, as well as participants' self-rating of awareness, were collectively at an average of 50% awareness level.

In terms of participants' expectations regarding data protection awareness, the findings showed an above-average level of 60% for the best age to start raising data protection awareness. However, a high frequency of participants reported low awareness levels. The data also revealed that although participants take awareness tips from organisational leads seriously, tips from data breach victims hold no significant value to them (6.7%), despite 29.7% of participants having suffered a data breach. These responses suggest that the initial POPIA training provided by the organisation may have influenced participants' opinions on data protection. Nonetheless, 99% of participants indicated that awareness training is necessary.

### C. Data Protection Awareness for Data Privacy

The participants' knowledge of the eight conditions of the POPIA was evaluated using a post-questionnaire. The findings indicated that the participants possessed a remarkable comprehension of the POPIA conditions. Their responses were consistently accurate, regardless of whether they were working remotely or from the office.

They achieved an average score of 95% on the assessment, implying a strong grasp of the POPIA conditions following their exposure to the data protection awareness artifact. Regarding the policies and procedures of the organisation, half of the participants agreed that they were clearly defined in terms of roles and responsibilities.

The responses to the post-questionnaire indicated that the participants had an average awareness of the organisation's policies and procedures. While most participants strongly agreed that hardware and software were securely configured, there may be a need for further awareness training regarding the implementation of the Bring-Your-Own-Device policy.

The research sub-questions aimed to explore how the employee's working environment(s) contributed to ensuring compliance with the POPIA by applying their data protection awareness. The survey established that most of the participants work from home, suggesting a higher rate of data collection, processing, storage, and sharing. However, those working from home have limited access to the tools they need to perform their duties compared to those working from the office, which may introduce a risk that employees may not apply their data protection knowledge consistently.

The findings indicated that most participants use their officially issued workstation or laptop, which is a securely configured device, to access the organisation's network. Most participants working from the office use the organisation's Wi-Fi instead of the LAN point to access the network. When working from home, employees are issued a data SIM, but the findings indicate that most of them use their personal Wi-Fi. Lastly, the findings show that only a few individuals have access to the work environment where participants process confidential information.

The study posits controlled access to an environment as a fundamental aspect of data protection. The data collected indicates that participants, both in their home and work offices, have the necessary controls in place for physical security. However, this finding raises concerns about the accuracy of participants' responses, as it is unlikely that those working from home have access doors, designated working areas, and password-protected printers. The findings demonstrate that the data protection awareness artifact was effective, well-designed, and had a positive impact on the participants. Additionally, the results reveal that the participants' basic understanding of data protection has increased. Overall, the information presented in this section supports the conclusion that the main research question has been satisfactorily addressed.

## 6. LIMITATIONS

While several organisations have undertaken similar initiatives to raise awareness on data protection, this study's scope is confined to a particular organisation and its workforce. Hence, the analysis presented in this research is delimited to a single organisation. Additionally, there were other limitations to this study. The research has identified several limitations that may affect the accuracy and validity of its findings. Firstly, there may be challenges with the research sample and selection process, which can introduce bias into the study. Secondly, the sample size of the study may be insufficient to identify significant relationships in the data, despite efforts to obtain a large sample.

Thirdly, by conducting an assessment of data protection awareness levels, it was found that there is a lack of previous research studies on the POPIA. This gap in knowledge limits the research's ability to build upon existing literature and make meaningful comparisons. To overcome this limitation, further research should focus on assessing data protection awareness levels within the context of the POPIA. This will contribute to a better understanding of data protection practices and inform the development of effective strategies to safeguard data.

The lack of previous research studies on the POPIA is a significant limitation for researchers who want to investigate the impact of the POPIA on businesses and organisations. One possible solution to this limitation is to conduct more research studies on POPIA. This would involve collecting and analyzing data on how POPIA affects different types of businesses and organisations, as well as how it affects individuals' personal information. Additionally, researchers could examine the challenges that businesses face in complying with the POPIA and how these challenges can be addressed. By conducting more research studies on POPIA, researchers can gain a deeper understanding of the impact of this Act and develop recommendations for how it can be improved to better protect individuals' personal information.

Furthermore, the technique or method used to collect data may not be optimal, which can hinder the thorough analysis of the research. The study recognizes that future research may benefit from the use of Machine Learning techniques to improve experiments and analysis. Finally, the study relies on organisational policies documentation, which may not reflect actual practices or provide a complete picture of the organisation's data protection measures. These limitations highlight the need for caution in generalising the findings of the study beyond the specific context of the organization and its participants.

### A. *Future Work*

The evaluation of employees' effectiveness in applying data protection awareness in a hybrid environment was the focus of this study. However, previous research by Baloyi and Kotze [43], Thorstensson [38], Abdullah and Hanifa [44], and Da Veiga et al. [45] have identified low consistency in the implementation of awareness training among employees, particularly in a hybrid working environment. Future research could investigate whether there has been an improvement in awareness, knowledge, and understanding of POPIA since its implementation. Additionally, to enhance data protection awareness for employees, the following research areas could be explored:

- The reasons for the slowdown in data protection training after the POPIA project implementation.
- How the maturity level of an organisation's cybersecurity contributes to its POPIA compliance.
- Extending the study to include the management team of the organisation.
- Examining the impact of employees' learning styles on awareness and how it affects the effectiveness of their application of data protection awareness.
- Applying Machine Learning techniques to predict the level of awareness among employees based on the collected data.

## 7. CONCLUSION

Evaluating the level of data protection awareness among employees presents a challenge for organisations. While achieving optimal effectiveness in this area may not be easy, identifying gaps and shortcomings is a good starting point. To achieve organisational goals, policies, standards, procedures, and business processes are crucial, and employees hold a significant responsibility in this regard. Thus, this research examined the influence of data protection awareness, considering individual employee preferences and circumstances that impact productivity. It is crucial to maintain sustained and ongoing refresher training and awareness initiatives to ensure the consistent and improved implementation of data protection measures. The study has developed a comprehensive framework for data protection awareness to enhance understanding of how to effectively promote data protection in both home-based and office-based work environments.

## REFERENCES

[1] R. Sayers, *Principles of awareness-raising for information literacy: A case study*. Communication and Information, UNESCO, 2006.

[2] O. Pantelic, K. Jovic, and S. Krstovic, "Cookies implementation analysis and the impact on user privacy regarding gdpr and ccpa regulations," *Sustainability*, vol. 14, no. 9, p. 5015, 2022.

[3] M. Nkongolo, J. P. van Deventer, and S. M. Kasongo, "Using deep packet inspection data to examine subscribers on the network," *Procedia Computer Science*, vol. 215, pp. 182–191, 2022.

[4] R. O. Mason, "Four ethical issues of the information age," *MIS quarterly*, pp. 5–12, 1986.

[5] K. Birch, D. Cochrane, and C. Ward, "Data as asset? the measurement, governance, and valuation of digital personal data by big tech," *Big Data & Society*, vol. 8, no. 1, p. 20539517211017308, 2021.

[6] M. Nkongolo, J. P. Van Deventer, and S. M. Kasongo, "Ugransome1819: A novel dataset for anomaly detection and zero-day threats," *Information*, vol. 12, no. 10, p. 405, 2021.

[7] M. Nkongolo, J. P. Van Deventer, S. M. Kasongo, S. R. Zahra, and J. Kipongo, "A cloud based optimization method for zero-day threats detection using genetic algorithm and ensemble learning," *Electronics*, vol. 11, no. 11, p. 1749, 2022.

[8] M. Nkongolo, J. P. van Deventer, and S. M. Kasongo, "The application of cyclostationary malware detection using boruta and pca," in *Computer Networks and Inventive Communication Technologies: Proceedings of Fifth ICCNCT 2022*. Springer, 2022, pp. 547–562.

[9] D. K. Citron, "Privacy injunctions," *Emory LJ*, vol. 71, p. 955, 2021.

[10] S. T. Margulis, "Conceptions of privacy: Current status and next steps," *Journal of Social Issues*, vol. 33, no. 3, pp. 5–21, 1977.

[11] R. Combley, *Cambridge business English dictionary*. Cambridge University Press, 2011.

[12] M. Nkongolo, J. P. van Deventer, S. M. Kasongo, and W. van der Walt, "Classifying social media using deep packet inspection data," in *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2022*. Springer, 2022, pp. 543–557.

[13] M. Nkongolo, "Using arima to predict the growth in the subscriber data usage," *Eng*, vol. 4, no. 1, pp. 92–120, 2023.

[14] M. Nkongolo, J. P. Van Deventer, S. M. Kasongo, W. Van Der Walt, R. Kalonji, and M. Pungwe, "Network policy enforcement: An intrusion prevention approach for critical infrastructures," in *2022 6th International Conference on Electronics, Communication and Aerospace Technology*. IEEE, 2022, pp. 686–692.

[15] R. FINN, "L., david wright, and michaeld., friedewald, m. 2013.'seven types of privacy'," *European Data Protection: Coming of Age, Springer, Dordrecht*, 2013.

[16] A. F. Westin, "Privacy and freedom," *Washington and Lee Law Review*, vol. 25, no. 1, p. 166, 1968.

[17] S. Baskarada and A. Koronios, "Data, information, knowledge, wisdom (dikw): a semiotic theoretical and empirical exploration of the hierarchy and its quality dimension," *Australasian Journal of Information Systems*, vol. 18, no. 1, 2013.

[18] A. Rossi, M. P. Arenas, E. Kocyigit, and M. Hani, "Challenges of protecting confidentiality in social media data and their ethical import," in *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2022, pp. 554–561.

[19] F. Belanger, J. S. Hiller, and W. J. Smith, "Trustworthiness in electronic commerce: the role of privacy, security, and site attributes," *The journal of strategic Information Systems*, vol. 11, no. 3-4, pp. 245–270, 2002.

[20] C. Doctorow, *Information doesn't want to be free: laws for the internet age*. McSweeney's, 2014.

[21] Y. Jordaan, "Public awareness and concerns relating to the protection of personal information act," 2015.

[22] G. G. Fuster, *The emergence of personal data protection as a fundamental right of the EU*. Springer Science & Business, 2014, vol. 16.

[23] J. Rubenfeld, "The right of privacy," *Harvard Law Review*, pp. 737–807, 1989.

[24] B. Van Niekerk, "An analysis of cyber-incidents in south africa," *The African Journal of Information and Communication*, vol. 20, pp. 113–132, 2017.

[25] H. Pieterse, "The cyber threat landscape in south africa: A 10-year review," *The African Journal of Information and Communication*, vol. 28, pp. 1–21, 2021.

[26] M. R. Mphatheni and W. Maluleke, "Cybersecurity as a response to combating cybercrime: Demystifying the prevailing threats and offering recommendations to the african regions," *International Journal of Research in Business and Social Science (2147-4478)*, vol. 11, no. 4, pp. 384–396, 2022.

[27] R. Gandhi, A. Sharma, W. Mahoney, W. Sousan, Q. Zhu, and P. Laplante, "Dimensions of cyber-attacks: Cultural, social, economic, and political," *IEEE Technology and Society Magazine*, vol. 30, no. 1, pp. 28–38, 2011.

[28] J. Warner, "Understanding cyber-crime in ghana: A view from below." *International journal of cyber criminology*, vol. 5, no. 1, 2011.

[29] O. E. Akinbowale, H. E. Klingelhöfer, and M. F. Zerihun, "The assessment of the impact of cyberfraud in the south african banking industry," *Journal of Financial Crime*, 2023.

[30] J. B. Rule and G. W. Greenleaf, *Global privacy protection: the first generation*. Edward Elgar Publishing, 2010.

[31] N. S. Netshakhuma, "An integrated approach to records management and information governance in south africa for sustainability," in *Intellectual Capital as a Precursor to Sustainable Corporate Social Responsibility*. IGI Global, 2023, pp. 36–50.

[32] L. I. Moraka and U. G. Singh, "The popia 7th condition framework for smes in gauteng," in *Computational Intelligence: Select Proceedings of InCITe 2022*. Springer, 2023, pp. 831–838.

[33] C. Walters, G. G. Mehl, P. Piraino, J. D. Jansen, and S. Kriger, "The impact of the pandemic-enforced lockdown on the scholarly productivity of women academics in south africa," *Research Policy*, vol. 51, no. 1, p. 104403, 2022.

[34] D. Banisar, "National comprehensive data protection/privacy laws and bills," *SSRN*, 2023. [Online]. Available: https://ssrn.com/abstract=1951416

[35] M. Grobler, J. Jansen van Vuuren, and J. Zaaiman, "Preparing south africa for cyber crime and cyber defense," 2013.

[36] R. Wilson, "Skills anticipation—the future of work and education," *International Journal of Educational Research*, vol. 61, pp. 101–110, 2013.

[37] L. Righetti and W. D. Smart, "The impact of robotics and automation on working conditions and employment [ethical, legal, and societal issues]," *Gene*, vol. 11, pp. 9afb–0c971f713d0c_story, 2021.

[38] E. Thorstensson, "The influence of working from home on employees' productivity: Comparative document analysis between the years 2000 and 2019-2020," 2020.

[39] L. Agostoni, "Remote working: Advices to reduce risks and boost productivity," *Iason Research Paper Series*, vol. 28, pp. 1–8, 2020.

[40] M. Saunders, P. Lewis, and A. Thornhill, *Research methods for business students*. Pearson education, 2009.

[41] R. M. Taylor, H. M. Colvin *et al.*, "Building a resilient workforce: Opportunities for the department of homeland security: Workshop summary," 2012.

[42] I. E. Allen and C. A. Seaman, "Likert scales and data analyses," *Quality progress*, vol. 40, no. 7, pp. 64–65, 2007.

[43] N. Baloyi and P. Kotzé, "Are organisations in south africa ready to comply with personal data protection or privacy legislation and regulations?" in *2017 IST-Africa Week Conference (IST-Africa)*. IEEE, 2017, pp. 1–11.

[44] H. Abdullah, "Towards the development of an information privacy protection awareness inititiative for data subjects and organizations," in *2021 National Computing Colleges Conference (NCCC)*. IEEE, 2021, pp. 1–7.

[45] A. da Veiga, E. Ochola, M. Mujinga, and E. Mwim, "Investigating data privacy evaluation criteria and requirements for e-commerce websites," in *Advanced Research in Technologies, Information, Innovation and Sustainability: Second International Conference, ARTIIS 2022, Santiago de Compostela, Spain, September 12–15, 2022, Revised Selected Papers, Part II*. Springer, 2022, pp. 297–307.