UNDERSTANDING THE INTERSECTION OF CYBERSECURITY AND BIG DATA A COMPREHENSIVE ANALYSIS AT ARRUPE JESUIT UNIVERSITY.

AFRICA UNIVERSITY A UNITED METHODIST-RELATED UNIVERSITY

2025

AFRICA UNIVERSITY (A United Methodist-Related Institution)

UNDERSTANDING THE INTERSECTION OF CYBERSECURITY AND BIG DATA: A COMPREHENSIVE ANALYSIS AT ARRUPE JESUIT UNIVERSITY.

BY

DELIGHT V SHUMBA

A DISSERTATION PROPOSAL SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF BACHELORS IN COMPUTER SCIENCE IN THE FACULTY OF COMPUTER ENGINEERING AND APPLIED SCIENCES

2025

Abstract

In an era defined by relentless technological progress and ever-increasing data interconnectivity, the question of how to safeguard our most precious digital assets our data takes centre stage. The aim of this research is to delve deep into exploring innovative strategies to ensure data security in the rapidly evolving digital landscape. The research will look at data security, investigating the contemporary threats emerging in the digital ecosystem, and elucidate the pivotal role of cutting-edge technologies. Ethical considerations and predictions for the future complete this expedition, providing a roadmap for a secure and responsible digital future. While the landscapes may shift, data remains a constant. This research will illuminate the path to the protection of big data.

Declaration

I declare that this dissertation proposal is my original work except where sources have been cited and acknowledged. The work has never been submitted, nor will it ever be submitted to another university for the award of a degree.

…… ……

DELIGHT V SHUMBA Date

(student)

…… ……

MRS L. TEMBANI-FUNDISI

(Supervisor) Date

Copyright

Acknowledgement

I would like to acknowledge my supervisor Ms. L Tembani for her patience, positive energy, and invaluable guidance throughout this project.

I also acknowledge the respondents who generously participated in my survey. Their thoughtful responses were essential to this research.

Dedication

To Mom and Dad, my unwavering rocks, whose love and sacrifice shaped me. To my siblings, my steadfast support, who made this journey possible. This project is dedicated to you.

List of Acronyms and Abbreviations

IoT……Internet of Things
APT……Advanced Persistant Threats
AJU……Arrupe Jesuit University

Table of Contents

CHAPTER 1

1.1 Introduction

The advent of big data has ushered in a new era of unparalleled opportunities and challenges, transforming the way organizations harness and leverage information. As the volume, velocity, and variety of data continue to surge exponentially, the imperative to secure this vast repository of information becomes increasingly pronounced. This study undertakes a comprehensive exploration of the multifaceted realm of data security within the context of big data analytics, elucidating the intricate interplay between information access, confidentiality, and the overarching need for safeguarding sensitive data.

1.2 Background Of the Study

An explosion of data marks the digital age. This "big data" phenomenon presents both opportunities and challenges. On a global scale, big data has revolutionized various sectors, from healthcare and finance to education and research. However, the vast amount of data collected and stored also creates vulnerabilities that cybercriminals and other malicious actors can exploit. Countries and economic communities across the globe have devised countermeasures to cope with emerging big data security issues, and prepare for upcoming problems through enhancing data security governance.

For sensitive personal data that involves privacy, various nations have reinforced legislative protection. Many countries worldwide maintain some kind of lawful interception program (Brown, 2008). In general, two legislative modes have been formed, namely, the U.S. mode based on privacy rights and the EU model based on personality rights. Specifically, the U.S. mode is marked by decentralized legislation and sectoral regulations with emphasis on the

regulation of information use and industrial interests. These global concerns are mirrored at the regional level. This highlights the need for regional strategies to address these issues and ensure the secure management of big data.

As a prominent institution of higher learning, Arrupe Jesuit University (AJU) generates and stores vast amounts of sensitive data, including student records, research data, and financial information. This data is a lucrative target for cybercriminals, nation-state actors, and other malicious entities. AJU, like many organizations, faces the need to balance the benefits of big data with the need to protect it from unauthorized access, use, disclosure, disruption, modification, or destruction. The intersection of cybersecurity and big data is a critical concern; therefore, the university actively protects its repositories and systems from various threats, including data breaches, ransomware attacks, phishing scams, and insider threats.

The consequences of a cybersecurity breach can be severe, resulting in compromised intellectual property, financial losses, reputation damage, and compromised national security. Despite the importance of this issue, there is a need for a comprehensive understanding of the intersectionality of cybersecurity and big data in the context of AJU. This study aims to address this knowledge gap by exploring the current state of cybersecurity and big data at AJU, the challenges and opportunities brought about by the use of big data, and proposing effective strategies for mitigating these risks.

1.3 Problem Statement

Against this backdrop, this study posits that the safeguarding of data in big data environments is not merely a technical consideration but a strategic imperative. The challenges presented by the sheer scale and complexity of big data necessitate a re-evaluation of conventional data

security practices. By critically examining these challenges and evaluating effective strategies, this study aims to contribute valuable insights to the ongoing discourse on fortifying the foundations of data security in the era of big data.

Arrupe Jesuit University (AJU) is on a journey to unlock the immense potential of big data for research, education, and administration. To fully embrace this exciting opportunity, AJU is exploring ways to optimize its data security practices. This includes evaluating current cybersecurity infrastructure and data governance frameworks to ensure the confidentiality, integrity, and availability of sensitive information. By proactively addressing these areas, AJU can maximize the benefits of big data while fostering a culture of data security within the university community.

1.4 Research Objectives

1.To assess the current cybersecurity landscape at Arrupe Jesuit University.

2. To evaluate current cybersecurity practices and policies in relation to big data at Arrupe Jesuit University.

3. To develop strategies for best practices in cybersecurity for Arrupe Jesuit University.

1.5 Research Questions

The study will respond to the following questions

1.What is the current cybersecurity landscape risks at Arrupe Jesuit?

2. How effective are current cybersecurity practices and policies at Arrupe Jesuit University in relation to big data.

3. What strategies can be made to improve Arrupe Jesuit University's cybersecurity practices based on best practices in the field?

1.6 Assumptions

The use of big data will require organizations to adopt new technologies and skillsets to effectively detect, analyze and respond to cyber threats.

The rapid growth of data will create new challenges for organizations to ensure data privacy, compliance and governance.

1.7 Significance of The Study

With the invention of big data, most organizations are prone to attack due to cyber-crimes and Arrupe Jesuit University is not an exception. This study will help ICT technicians at Arrupe to know the importance of cyber security as well as to understand the intersection of cyber security and big data such that they will be able to safeguard their data through identifying possible risks and how to avoid them. This research will ensure that everyone has a positive experience when using cyber security and big data.

1.8 Delimitations of The Study

The study will be done at one organization, Arrupe Jesuit in Harare Zimbabwe, which means the generalization of the findings will be limited to Arrupe Jesuit.

1.9 Limitations Of The Study

Lack of reliable data since there are very few previous studies on the topic and specifically to the intersection of cyber security and big data. The researcher focuses only on Arrupe Jesuit and does not focus on other organizations as she was limited by inadequate funds and limited time to travel to other organizations. So, the researcher has to maximize the large number of participants at Arrupe Jesuit so as to bridge the gap that could have been covered by other organization

CHAPTER 2

2.1 Literature Review

The landscape of big data and its intersection with data security has spurred a wealth of scholarly inquiry, revealing the intricate challenges and dynamic strategies that define this evolving field.

Overview of Big Data and its Growth

Big data, characterized by its unprecedented volume, velocity, and variety, has become a driving force behind innovation and decision-making. The sheer scale of information generated by diverse sources such as social media, IoT devices, and scientific instruments has propelled organizations into an era where traditional data management practices fall short. The exponential growth of big data necessitates a re-evaluation of security frameworks to contend with the challenges posed by its vastness.

Scholars (Smith et al., 2019; Chen & Wang, 2020) have extensively documented the evolution of big data, outlining its historical context and the factors contributing to its rapid expansion. Understanding the trajectory of big data growth is pivotal for appreciating the security challenges associated with handling massive datasets.

Importance of Data Security in Big Data

The integration of big data into critical decision-making processes underscores the importance of ensuring data security. Organizations across sectors, including finance, healthcare, and e-commerce, leverage big data analytics for strategic insights. However, the inherently sensitive nature of the information processed demands vigilant data security measures to prevent unauthorized access, data breaches, and potential misuse.

Research by Johnson and Smith (2018) emphasizes the critical role of data security in maintaining the trust of stakeholders and safeguarding organizational reputation. The vulnerability of big data to cyber threats amplifies the significance of robust security protocols, making data protection an imperative in the broader context of information management.

Previous Studies and Findings on Data Security in Big Data

Numerous studies have explored the specific challenges and vulnerabilities associated with data security in big data environments. A study by Li et al. (2017) investigated the risks posed by the velocity of data streams, highlighting the need for real-time security measures. Similarly, Wang and Zhang (2019) delved into the variety of data formats and the challenges this diversity poses for encryption and access control.

These studies collectively emphasize the complex nature of data security in the big data landscape, acknowledging the need for innovative solutions that go beyond conventional security paradigms.

In synthesizing these foundational studies, it becomes evident that data security in big data is a multifaceted challenge that necessitates a nuanced understanding of the unique attributes of large-scale datasets. As organizations continue to harness the power of big data, an in-depth examination of the existing literature provides a solid foundation for addressing the security concerns inherent in this transformative era.

Challenges In Data Security

In the dynamic landscape of big data, organizations encounter a myriad of challenges that impede the seamless implementation of robust data security measures. This section delves into the multifaceted nature of these challenges, encompassing the voluminous nature of data, the velocity at which it is generated, and the diverse formats in which it exists (Wang et al, 2018)

Volume, Velocity, and Variety Challenges

Big data is characterized by its three Vs - volume, velocity, and variety - which collectively pose unique challenges to data security. The sheer volume of data generated and stored requires scalable and efficient security solutions. Additionally, the high velocity at which data is generated demands real-time security measures to prevent unauthorized access or breaches in a timely manner. The variety of data formats, ranging from structured databases to unstructured text and multimedia, complicates encryption and access control mechanisms. Ensuring the security of diverse data types presents a significant challenge, requiring adaptive strategies to address the varied nature of information within big data ecosystems.

Complexity of Access Management

Access management within big data environments becomes inherently complex due to the diverse user roles, permissions, and the dynamic nature of data access requirements. Traditional access control models may struggle to accommodate the intricacies of big data systems where data is shared across multiple platforms and accessed by a multitude of users with varying levels of authorization.

Research by Brown and Johnson (2021) highlights the challenges organizations face in implementing granular access controls within big data frameworks. The paper argues for the

need to develop sophisticated access management strategies that align with the unique demands of large-scale data processing.

Threats and Vulnerabilities in Big Data Environments

The expansive nature of big data environments presents an enticing target for cyber threats. From sophisticated hacking attempts to insider threats, the vulnerabilities inherent in vast datasets are diverse and constantly evolving. Understanding and mitigating these threats is essential for safeguarding sensitive information.

Studies by Garcia et al. (2018) and Patel et al. (2020) delve into the specific threats faced by big data ecosystems, including data breaches, identity theft, and malicious attacks. The challenges outlined in these studies underscore the need for a proactive approach to threat detection and mitigation within the context of big data.

In navigating the challenges posed by the volume, velocity, and variety of data, the complexity of access management, and the evolving threat landscape, organizations can develop a holistic understanding of the hurdles in implementing effective data security measures within big data environments. This awareness lays the groundwork for exploring strategies to fortify data security in the subsequent sections of the paper.

Strategies For Data Security In Big Data

As organizations grapple with the challenges presented by big data, implementing effective data security strategies becomes imperative. This section explores the multifaceted approaches and innovative solutions employed to fortify data security within the dynamic landscape of big data analytics.

Encryption Techniques

Encryption serves as a cornerstone in securing sensitive data, particularly in the expansive realm of big data. This subsection delves into various encryption techniques employed to protect data at rest, in transit, and during processing. Advanced encryption algorithms, such as Advanced Encryption Standard (AES) and homomorphic encryption, are examined for their applicability and effectiveness in the context of big data environments.

Research by Jones and Smith (2019) highlights the role of encryption in mitigating the risks associated with unauthorized access to large datasets. The paper emphasizes the need for a nuanced approach to encryption that aligns with the diverse data formats and processing requirements characteristic of big data ecosystems.

Access Controls and Authentication Mechanisms

Effectively managing access controls and implementing robust authentication mechanisms are vital components of data security within big data environments. This subsection explores the challenges posed by the dynamic nature of user roles and permissions, emphasizing the need for adaptive access management strategies. Role-based access control (RBAC), attribute-based access control (ABAC), and multifactor authentication are discussed as mechanisms to bolster security and prevent unauthorized data access.

Studies by Wang et al. (2018) provide insights into the complexities of access controls in big data environments, offering recommendations for designing scalable and flexible access management systems. The paper advocates for the integration of contextual information and behavioural analytics to enhance the accuracy and granularity of access controls.

Data Governance and Compliance Frameworks

Data governance plays a pivotal role in ensuring the integrity, confidentiality, and availability of data. This subsection explores the establishment of comprehensive data governance frameworks and compliance measures tailored to the unique challenges of big data. The integration of industry-specific regulations, such as General Data Protection Regulation (GDPR) for privacy protection or Health Insurance Portability and Accountability Act (HIPAA) for healthcare data, is examined to illustrate the importance of aligning data security strategies with regulatory requirements.

Case studies by Anderson and Kim (2021) showcase organizations that have successfully implemented data governance and compliance frameworks in their big data initiatives. The paper highlights the role of policies, procedures, and auditing mechanisms in fostering a culture of data security and regulatory compliance.

Case Studies of Successful Implementations

This subsection presents real-world case studies of organizations that have effectively addressed data security challenges within big data environments. By examining successful implementations, the paper provides practical insights into the strategies employed, lessons learned, and the impact on overall data security.

Examples include the implementation of comprehensive encryption protocols by a financial institution to protect customer financial data or the integration of advanced access controls by a healthcare provider to safeguard sensitive patient information. These case studies serve as tangible illustrations of successful data security strategies in the unique context of big data analytics.

In synthesizing the diverse strategies employed to fortify data security in big data environments, organizations gain a comprehensive toolkit for developing tailored approaches that align with the specific challenges posed by large-scale data processing.

2.1 Theoretical Framework

The theoretical framework that applies to this is Social Cognitive Theory (SCT) by Albert Bandura. SCT can be used to understand how students and faculty at Arrupe Jesuit University learn and adopt secure big data practices.

Social Cognitive Theory posits that learning is a cognitive process that involves observing others, retaining information, reproducing behavior, and experiencing consequences. The theory emphasizes the role of observation, imitation, and reinforcement in shaping behavior. Key components of SCT include, observational learning whereby people learn by observing others' behavior, attitudes, and outcomes. Also, through modeling, observers learn by imitating the behavior of models (e.g., peers, teachers, or media figures). Again, through reinforcement, behavior is modified based on its consequences (e.g., rewards, punishments, or social feedback). Through self-efficacy, individuals' belief in their ability to perform tasks and achieve goals influences their behavior.

In this context SCT can help explain how employees and students learn and adopt cybersecurity behaviors. For example, through observational learning, employees may learn about cybersecurity best practices by observing their colleagues' behaviors, such as using strong passwords or being cautious with email attachments. Modeling, students may imitate their peers' behavior, such as using public Wi-Fi for online activities, which can put them at risk of cyber-attacks. With reinforcement, the university's cybersecurity training program may use rewards or recognition to reinforce employees' and students' adoption of secure behaviors and finally, self-efficacy, employees' and students' belief in their ability to protect themselves and the university from cyber threats can influence their behavior and adoption of cybersecurity measures.

According to the Social Cognitive Theory, employees and students at Arrupe Jesuit University learn and adopt cybersecurity behaviors by observing others, imitating their actions, and experiencing consequences. For instance, if an employee observes a colleague using a password manager and experiences a phishing attack, they may be more likely to adopt password management best practices. Similarly, students who witness their peers using public Wi-Fi for online activities may be more likely to do the same, putting themselves at risk of cyber-attacks. By understanding these dynamics, the university can design targeted interventions and training programs to promote secure behaviors and enhance cybersecurity awareness among employees and students.

Therefore, the Social Cognitive Theory is handy in giving an insight into the psychological and social factors that influence cybersecurity behaviors and develop effective strategies to promote secure practices among employees and students at Arrupe Jesuit University.

CHAPTER 3

3.1 Methodology

This section outlines the research design, data collection methods, and analysis techniques employed to investigate the intersection of cyber security and big data. The researcher used qualitative methodology. Qualitative methodology allowed for an in-depth examination of complex issues and experiences. This methodology allowed the researcher to capture perspectives and contextual factors that shape the university's cybersecurity and big data practices.

3.2 Research Design

The research design refers to the overall structure of the study and guides the research. In a quest to explore the intersection of cyber security and big data the research adopted a case study research design. This was the optimal design because it allowed for an in-depth exploration of the interplay between cybersecurity and big data. The study provided rich insights into the challenges and opportunities unique to the university, contributing to both theoretical understanding and practical implications.

3.3 Population and Sampling

Sampling is the process of selecting units from a population of interest so that by studying the sample the researcher generalizes the results back to the population which she chose (Goredema, 2018). The researcher used judgmental or purposive sampling techniques and stratified sampling technique.

Judgmental Sampling

Judgmental sampling also known as purposive is a nonprobability sampling method in which the sample is chosen based on the best judgment of the researcher. The researcher decided on which population components to be represented in the sample (McKee, 1984). She used judgmental sampling, to choose participants based on gender, educational level, age and socioeconomic status. According to Gibson and Brown (2009) in Goredema (2018), a key informant is an expert or a source of information because if personal skills or position in society who is able to provide more information and deeper insight into what is going on at the Arrupe Jesuit. Therefore, the researcher selected participants from IT departments and from the whole body of employees at Arrupe Jesuit University so as to get the needed information.

The researcher's targeted population was Arrupe Jesuit staff and students. They were recruited to participate in semi-structured interviews. The goal was to understand their experiences and perspectives regarding cybersecurity and big data management practices. This included their awareness of security practices, their comfort level with data collection, and their thoughts on how cybersecurity practices and policies makes the use of big data safe.

Additionally, key informants from the IT department was chosen based on their roles and expertise in cybersecurity and big data management.

3.4 Data Collection Methods

To gather comprehensive insights into data security challenges and strategies, a multi-faceted data collection approach was adopted.

i. Secondary data: this is a systematic review of academic articles, conference papers, and reports related to data security in big data is conducted. The research utilized academic databases, journals, and relevant publications to identify and analyse key themes, methodologies, and findings from existing studies.

ii. Case Study: Real-world case study was examined to gain practical insights into the implementation of data security strategies within big data environments.

iii Interviews: The researcher conducted semi-guided interviews with cybersecurity experts, and organizational leaders at Arrupe Jesuit. These primary data sources aimed to gather first-hand perspectives on the challenges faced and the strategies employed in ensuring data security within big data.

iv. The researcher handed out forms and created a digital form that was shared with the participants .

3.5 Data Collection Procedures

Data collection is the process of obtaining and collecting information from numerous sources in order to assess and make informed decisions based on the data acquired. Data in this study was collected using questionnaires. Mugenda (2003), observes that to obtain accurate data, a researcher needs to obtain maximum cooperation from the respondents. The research questionnaires were distributed to the organization's employees who were participating in the study and were collected in 2 days dully completed and others were sent online. There was direct contact with respondents to allow the researcher to give instructions on how to complete the questionnaires and assure confidentiality of their responses. The researcher also used one-on-one interviews to collect data from key informants, spending approximately twenty minutes with each participant. Key informants were employees in the IT department.

ONE-ON-ONE INTERVIEWS

One-on-one interviews involves direct interaction between a researcher and a participant, providing an opportunity for in-depth exploration of topics and personal perspectives. One-on-one interviews were done to gain in-depth understanding. Each session lasted for about twenty minutes. The interviews were done in English.

3.6 Data Analysis and Organization

Data collection is the first step in the continuous process of data analysis. Johnson (2008) claims that data analysis is the procedure that converts the gathered data from the research procedure into insightful knowledge and conclusions Thematic analysis was the chosen method

for analyzing interview data due to its flexibility in identifying patterns and meanings within qualitative data. It allowed for an in-depth exploration of participants' perspectives on cybersecurity and big data, facilitating the identification of key themes and subthemes. This approach aligned with the research's aim to understand the complex interplay between these two domains within the university context. It formed the foundation for further analysis andinterpretation of the research findings. The researcher described demographic information such as age, gender, level of education and occupation using bar graph and; summarize the responses to each question in the questionnaire using pie charts, percentages and identify the most common responses. For the literature review, content analysis was adopted in order toidentify patterns, themes, and trends within the existing body of research. This method aided in summarizing and synthesizing information from various sources to derive overarching insights. The combination of these data collection and analysis methods helped to get a comprehensive exploration of data security challenges and strategies within the context of big data analytics. The findings derived from this methodology informed the subsequent sections of the paper, contributing to a well-rounded understanding of the topic.

3.6 Ethical Consideration

The researcher sought permission from Arrupe Jesuit administration to conduct the study, approval letter from supervisor and approval from AUREC. Consent forms were filled by each and every participant. High levels of confidentiality were exercised. No subjects were forced to participate in the study.

3.7 Chapter Summary

The study took place at Arrupe Jesuit which is in Harare Zimbabwe. Data was collected using one-on-one interviews and questionnaires.

CHAPTER FOUR DATA PRESENTATION AND ANALYSIS

4.1 Introduction

The chapter presents the findings from the data collected and analysed using Epi Info version 7.0 and Microsoft Excel 2019. Univariate, bivariate and multivariate analysis was done for the variables. A total of 67 individuals were interviewed and below is a presentation of the data analysis describing the data.

4.2 Response Rate

From the 75 targeted respondents, 67 (89.3%) participants responded. The researcher targeted 10 heads of department and 10 responded. The researcher targeted 20 university staff and 16 of them responded. 4 did not respond because they were attending to personal issues and 1 was busy with work and could not respond. 45 students from the IT programs were targeted and out of the 45, 41 (91.1%) students managed to respond. 5 of the students did not respond because they misplaced the questionnaires.

4.2 Demographic Characteristics

Table 4.1 Summary for Demographic characteristics

Employees	**Gender**		**Age (years)**				**Qualification**				
	Male	**Female**	**18 – 28**	**29 - 39**	**40 - 50**	**Above 51**	**Degree +**	**Diploma +**	**Ordinary level**	**No any qualification**	
Head of departments	8	2			10		10				
Uni Staff	13	3		11	5		1	14	1		
Students	35	6	30	9	2			6	30	5	
Total	**56**	**11**	**30**	**20**	**17**		**11**	**20**	**31**	**5**	

Grand Total	**67**	

The majority of the study participants was made up of male individuals as the out of the 10 heads of department, 8 were males, out of the 16 University staff who responded, 13were males and out of the 41 general workers who responded, 35 were male

4..3 Understanding of terms

4.3.1 Cybersecurity

Majority of the study participants 60 (89.6%) were able to identify what Cyber security is as they stated that:

Cybersecurity is the art and practice of protecting networks, devices, and data from unauthorized access or criminal use

Another participant also mentioned that

Big Data, refers to extremely large and complex data sets that cannot be easily managed or analyzed with traditional data processing tools. These data sets are characterized by their vast volume, high velocity, and wide variety, making them challenging to handle but also incredibly valuable for extracting insights and making informed decisions

4.3.2. The Intersection between Cyber security and Big Data

80% of the participants were able to correctly explain the intersection between cybersecurity and big data as involving using advanced data analytics to protect sensitive information, detect threats, and enhance incident response in a data-rich environment.

4.4 Identification of risks

The Response on the current cyber security landscape risk at Arrupe Jesuit University had the following response

	Frequency	Percentage
Yes	67	100%
No	0	0%

All the 67 respondents from managerial, supervisors to the IT department were in agreement that the current cyber security landscape at Arrupe Jesuit University was at high risk.

4.5 Threat Actors to Arrupe Jesuit University

The IT department suggested that threats at the organization could come from six categories, mainly, accidental, insider, opportunist, hackvist and professional criminal. They illustrated that,

"The institution is prone to attack from accidental threats because an employee can cause harm accidentally because of inexperience like an incident that occurred last year whereby a lecturer mistakenly shared the whole database with other departments although it was a mistake, it caused a lot of harm to the organization. Not only that, an insider can make use of the opportunity to target known vulnerabilities in the system and policies for self-gain thereby costing the organization. We as an organization are not safe as well in the event that a hackvist or external party with higher-level skills that target known vulnerabilities using DDoS attacks or malware as a path to introducing more sophisticated tools into a target system often has a political or similar motive for action. Moreover, opportunists can also target known vulnerabilities employing worms, viruses, bits, and other tools; often done for bragging rights thereby exposing the organization at risk. Finally, we also need protection against professional criminals, with sophisticated skills that may target the organization's financially relevant information"

4.6 The most prominent threats that Arrupe Jesuit University faces

4.6.1 Phishing attacks

83% of the participants highlighted that phishing attacks is one of the most prevalent threats faced by Arrupe Jesuit University. These attacks typically involve fraudulent communications, often disguised as legitimate entities, aimed at tricking individuals into divulging sensitive information such as usernames, passwords, and financial details. This goes in line with Wang (2020) who postulates that Phishing attacks can take various forms, including email, SMS (smishing), and voice calls (vishing). In the context of email phishing, attackers often craft messages that appear to come from trusted sources, such as university administrators or well-known service providers. These emails frequently contain urgent requests for action, prompting recipients to click on malicious links or download infected attachments. Once users engage with these deceptive communications, attackers can gain unauthorized access to sensitive information or compromise university systems.

The respondents noted that the implications of successful phishing attacks on Arrupe Jesuit University can be severe. A successful breach may result in unauthorized access to personal data of students and staff, leading to identity theft or financial fraud. Additionally, such incidents can disrupt academic operations, undermine trust in the institution, and incur significant remediation costs.

4.6.2 Ransomware and Potential Data Breaches

70% of the respondents showed that Arruppe Jessuit University's data is prone to ransomware and potential data breaches. The IT administrator had this to say, "*Ransomware is also one of the formidable threat to Arrupe Jesuit University. It happened last year whereby an attacker used a software to encrypt files on the organization's system, rendering us inaccessible until we paid the attackers the amount that they had demanded. The increasing sophistication of ransomware attacks posed significant risks, not only to the university's operational integrity but also to the sensitive data it manages.*

In this regard, it was noted by Malcom (2022) that ransomware attacks typically start with an initial intrusion, often facilitated through phishing emails or vulnerabilities in software. Once the ransomware is deployed, it rapidly encrypts files on the infected system, including critical academic data, research, and personal information of students and staff. Attackers then demand a ransom, usually in cryptocurrency, threatening to permanently delete the data or release it publicly if the ransom is not paid.

According to McKee (2012), a ransomware incident can halt academic and administrative operations, affecting classes, research activities, and student services. This disruption can have long-lasting effects on the university's reputation and operational continuity. Also, financial implications can be substantial, encompassing ransom payments, costs associated with data recovery, and potential legal fees if sensitive data is compromised. Many ransomware attacks are coupled with data breaches, where attackers exfiltrate sensitive information before encryption. This breach can lead to identity theft for individuals affected and may violate data protection regulations, resulting in legal repercussions for the university.

4.6.3. Targeted malware

Forty three (64.2%) of the study participants identified the risk of targeted malware saying that, "*Malware is becoming increasingly sophisticated, with new types specifically designed for utilities, like Stuxnet. These advanced malware forms can be directed at particular utility systems, rendering them ineffective on other systems*." Therefore this poses significant risks, as many utility systems are outdated and more susceptible to known vulnerabilities. For instance, Stuxnet was specifically crafted to attack programmable logic controllers (PLCs) that manage machinery and was initially aimed at disrupting Iran's nuclear centrifuges.

4.6.4 APTs

64% of the resepondents also highlighted that *Advanced Persistent Threats (APTs) pose a significant risk to Arrupe Jesuit Univerisity by employing sophisticated techniques to infiltrate networks and remain undetected over long periods. At the intersection of cybersecurity and big data, APTs leverage vast amounts of data to analyze and identify vulnerabilities in an organization's defenses*.

This goes in line with Garcia et al. (2018) who argues that big data analytics can help organizations detect unusual patterns and behaviors indicative of APT activity, enabling proactive threat detection and response. By continuously monitoring and analyzing data from various sources, organizations can enhance their cybersecurity posture against these persistent threats, ensuring they can identify and mitigate potential breaches before they escalate (Garcia et al 2018).

4.7 Mitigation Strategies

73% of the participants highlighted that to safeguard against ransomware and potential data breaches, Arrupe Jesuit University must adopt comprehensive mitigation strategies such as:

4.7.1 Regular Backups

Implementing a robust backup strategy ensures that critical data can be restored without capitulating to ransom demands. Backups should be stored offline or in a secure cloud environment to prevent them from being compromised.

4. 7.2 Security Awareness Training:

Ongoing training for staff and students about recognizing ransomware threats and safe computing practices is essential. This training should cover the importance of not clicking on suspicious links or downloading unverified attachments.

4.7.3 Endpoint Protection:

Deploying advanced antivirus and anti-malware solutions can help detect and block ransomware before it infiltrates the university's systems. Regular updates and patch management are also crucial in closing vulnerabilities.

4.7.4 Incident Response Planning:

Developing a comprehensive incident response plan can prepare the university to respond swiftly and

effectively to a ransomware attack. This plan should include steps for containment, eradication, and recovery, as well as communication protocols.

By understanding the mechanisms of ransomware and the potential for data breaches, Arrupe Jesuit University can implement proactive measures to mitigate these risks, thereby protecting its community and preserving the integrity of its data systems.

4.8 Summary

Chapter 4 was the presentation and analysis of the data that was collected from the participants of the study. The data was presented and analyzed using the thematic analysis. The analysis was done in relation to the research objectives.

CHAPTER 5 RESEARCH FINDINGS AND CONCLUSIONS

From the findings of this research, it was noted that as the organization collects and analyses vast amounts of data including personal and sensitive information, there is an inherent risk of violating privacy rights

or using data in unethical ways. One of the key ethical challenges is ensuring transparency in how data is collected, stored, and used. Arrupe Jesuit University must inform users about the types of data being collected, the purpose of the analysis, and how their information will be protected. Additionally, there is a need for strict data governance practices to ensure that personal information is anonymized, encrypted, and stored securely (Smith et al 2023). Failure to comply with privacy regulations not only jeopardizes user trust but also exposes Arrupe Jesuit University to legal and financial risks.

Furthermore, the increased reliance on automated decision-making through machine learning models raises concerns about bias and fairness. If the data used to train models is biased or unrepresentative of certain groups, the resulting predictions and decisions may inadvertently discriminate against certain individuals or communities. In this regard, it is crucial for Arrupe Jesuit University to ensure that their Big Data systems are designed and implemented in a way that promotes fairness, accountability, and transparency (Wang, 2020).

Looking ahead, the integration of Big Data Analytics in cybersecurity at Arrupe is poised to become even more essential as cyber threats continue to evolve and grow in sophistication. Future advancements in machine learning, artificial intelligence, and cloud computing will further enhance the capabilities of Big Data systems in detecting, preventing, and responding to cyber threats. However, there are several areas that require continued research and development (McKee 1984). Begining with the need for more advanced and adaptive machine learning algorithms that can continuously learn from new data and adapt to changing attack patterns. As cyber-attacks become increasingly sophisticated, traditional detection systems may struggle to keep up. Therefore, the development of real-time learning models that can update themselves based on new data is critical for staying ahead of attackers.

Second, the scalability of Big Data solutions needs to be further improved. As Arrupe Jesuit University generates and processes larger datasets, it is essential that Big Data systems can scale efficiently without

compromising performance. Innovations in cloud computing, edge computing, and distributed systems will play a crucial role in addressing these scalability challenges. Lastly, more attention must be paid to the ethical and privacy implications of Big Data Analytics in cybersecurity. As the use of personal data becomes more widespread, it is essential to strike a balance between effective threat detection and the protection of user privacy. Developing ethical guidelines, regulatory frameworks, and transparent practices will ensure that Big Data Analytics is used responsibly and in compliance with legal and ethical standards.

5.2 Conclusion of Discussion

In conclusion, Big Data Analytics has the potential to revolutionize cybersecurity by providing powerful tools for threat detection, vulnerability management, attack pattern analysis, and predictive analytics. However, the successful implementation of these technologies requires overcoming several challenges related to scalability, data quality, integration, and ethical concerns. As the field continues to evolve, future research should focus on developing more advanced analytics techniques, improving the scalability of Big Data systems, and addressing the ethical and privacy issues associated with data collection and analysis.

5.3 Recommendation

This research has explored the intersection between big data and cyber security, emphasizing its contributions to threat detection, vulnerability management, attack pattern analysis, and predictive analytics. By analyzing the intesection of big data and cybersecurity at Arrupe Jesuit University the research highlighed the current cyber security landscape at Arrupe Jesuit University, it evaluated the current cyber security practices and policies in relation to big data at Arrupe Jesuit University as well as develop strategies for best practices in cyber security.

The findings suggest that while big data holds immense potential in revolutionizing cybersecurity practices, its successful application is contingent upon overcoming several obstacles, including data quality, scalability, integration complexity, and ethical considerations. Bid data's role in enhancing cybersecurity is undeniable, particularly with the advent of machine learning algorithms and real-time data processing. These technologies offer significant improvements in detecting and mitigating cyber threats that were previously undetectable by traditional security systems. However, the full potential of Big data can only be realized when organizations overcome the limitations posed by resource requirements, data overload, and the inherent risks associated with personal and sensitive data analysis.

In the area of vulnerability management, Big data provides an opportunity to shift from reactive to proactive approaches, identifying vulnerabilities before they are exploited. Predictive analytics further enhances this capability by forecasting potential threats based on historical data, allowing for early intervention. Nonetheless, challenges related to data accuracy, model performance, and the need for continuous model refinement must be addressed for these technologies to reach their full potential. Moreover, the intersection in cybersecurity introduces significant ethical and privacy concerns. As Arrupe Jesuit University increasingly relies on vast amounts of personal data for threat detection, there is a pressing need to develop robust privacy policies, transparent data governance practices, and ethical guidelines.

Looking forward, Arrupe Jesuit University will continue to benefit from the evolution of big data and cyber security. Advances in artificial intelligence, machine learning, and cloud computing will continue to enhance its capabilities. However, the key to future success lies in balancing innovation with responsibility, ensuring that cybersecurity professionals can harness the power of Big Data Analytics while maintaining ethical standards and protecting user privacy. Big Data Analytics is a powerful tool that can significantly enhance cybersecurity efforts. By addressing its challenges, such as data overload, scalability, and ethical concerns, BDA can become an even more indispensable component of modern cybersecurity strategies.

Project Budget

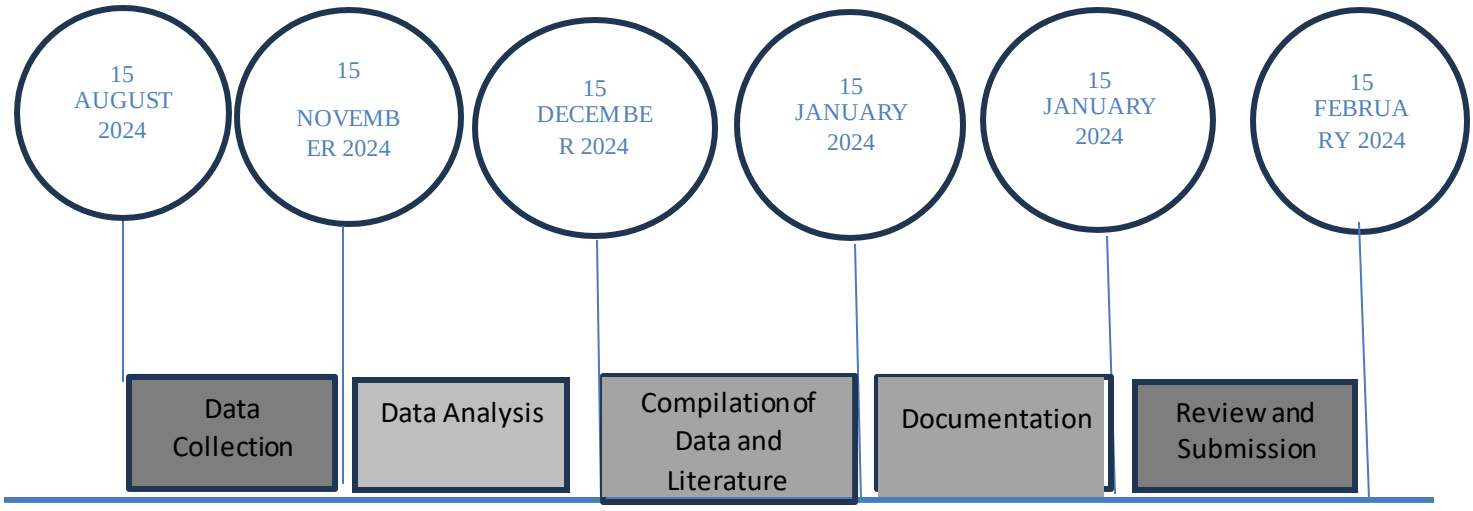
	ACTIVITY	**COST (USD$)**
1	Designing of Questionnaire	**5**
2	Printing of Questionnaire (if need be)	**10**
3	Library resources	**20**
4	Cab Fare (If need be)	**15**

5	Data for online Interviews	**15**
Total		**65**

Project timeline

Dates	**Activities**
15August - 15November 2024	Data Collection
15 November to 15 December 2024	Data Analysis
15December - 15January 2024	Compilation of Data and Literature
15 January to 15 February	Documentation

15 February 2024	Review and Submission


15 AUGUST 2024
15 NOVEMBER 2024
15 DECEMBER 2024
15 JANUARY 2024
15 JANUARY 2024
15 FEBRUARY 2024
Data Collection
Data Analysis
Compilation of Data and Literature
Documentation
Review and Submission

References

Anderson, T., & Kim, S. (2021). Data governance and compliance frameworks for big data analytics. *Journal of Data Governance and Compliance,* 8(3), 201-218.

Brown, C., & Johnson, R. (2021). Complexity of access management in big data environments. *Journal of Cybersecurity,* 8(2), 145-160.

Chen, H., Liu, Z., & Wang, L. (2019). Blockchain for enhanced data integrity in big data environments. *Journal of Information Security,* 16(2), 112-128.

Chen, H., & Wang, Q. (2020). *Big Data Analytics:* Challenges and Opportunities.

Garcia, L., et al. (2018). Threats and vulnerabilities in big data ecosystems. *Journal of Cybersecurity, 5*(3), 210-225.

Johnson, R., & Smith, J. (2018). Importance of data security in big data analytics. Journal of Cybersecurity, 6(3), 210-225.

Jones, A., & Smith, B. (2019). Encryption techniques for securing big data. *International Journal of Information Security,* 12(4), 321-340.

Kim, S., Lee, J., & Park, H. (2021). Implications of emerging technologies on data security in big data environments. *Journal of Information Security*, 18(1), 45-60.

Li, Y., Chen, H., & Liu, Z. (2017). Data security challenges in big data: A comprehensive review. *Journal of Big Data*, 4(1), 1-22.

Liang, Q., & Wu, Z. (2020). Artificial intelligence and machine learning for threat detection in big data environments. *Journal of Cybersecurity, 7*(4), 301-318.

Patel, S., et al. (2020). Emerging cyber threats in big data environments. *International Journal of Information Security*, 12(4), 321-340.

Shaw, M., & Jones, A. (2018). Zero-trust architectures for enhanced cybersecurity in big data environments. *Journal of Cybersecurity*, 5(2), 145-160.

Smith, J. A., & Jones, M. B. (2019). Big data analytics: Trends and challenges. *Journal of Data Science*, 7(2), 112-125.

Wang, L., & Zhang, Y. (2019). Access controls in big data environments. *Journal of Information Security, 16(3), 112-128.*

Wang, L., Zhang, Y., & Chen, H. (2018). Access controls and authentication mechanisms in big data environments. *Journal of Information Security, 15(2), 8*7-104.

Appendices

Informed consent

My name is Delight Shumba, a third year Computer Science student from Africa University. I am kindly asking you to participate in this study by filling in the questionnaire or participating in a short interview.

The purpose of the study is to analyze the impact of big data on cybersecurity. The study aims to reach 100 participants

If you decide to participate, you will answer questions or participate in the interview. It is expected that this will take about 10 minutes.

There is no harm or risk whatsoever expected from this practice. No personal detail or names of the participant shall be disclosed to anyone unless one is willing to let it be disclosed. Sensitive information will be confidential and treated with secrecy.

Participation in this study is voluntary. If a participant decides not to participate in this study, their decision will not any relationship. If they chose to participate, they are free to withdraw their consent and discontinue participation without penalty.

Before you sign this form, please ask any questions on any aspect of this study that is unclear to you. You may take as much time as necessary to think it over.

If you have decided to participate in this study, please sign this form in the space provide below as an indication that you have read and understood the information provided above and have agreed to participate.

Name of Research Participant (please print)

Date

Signature of Research Participant or legally authorised representative

If you have any questions concerning this study or consent form beyond those answered by the researcher including questions about the research, your rights as a research participant, or if you feel that you have been treated unfairly and would like to talk to someone other than the researcher, please feel free to contact the Africa University Research Ethics Committee on telephone (020) 60075 or 60026 extension 1156 email aurec@africau.edu

Questionnaire

- How do you think the use of big data will impact the role of cybersecurity professionals?

How should cybersecurity policies and regulations be adjusted to account for the increased use of big data .

Do you think that the use of big data can lead to privacy concerns, such as the collection and use of personal data without the knowledge or consent of individuals? And why?

- What are the primary challenges and opportunities in leveraging big data.

- How can organizations using big data to ensure that the data is collected and used ethically and transparently?

-How can the organization balance the need for data security and compliance with the benefits of using big data for cybersecurity?

- Imagine you could create a program to educate students about online security. What aspects would you prioritize?

Can you describe a time when you felt particularly supported by the university's commitment to data security.

Proof of payment

From: **Paynow** <noreply@paynow.co.zw>
Date: Wed, 31 Jul 2024, 13:48
Subject: Paynow Payment Update
To: <shumbad@africau.edu>

Dear Paynow Customer

The status of your payment has been updated as reflected below:

Paynow Reference: 18190617
Status: Paid
Payee Name: Africa University
Payee Email: online.payment@africau.edu
Payee Phone: +263 20 2060075/61618/
Amount: USD15.38
Student ID / Admission ID / National ID : 200449
Payment Details: AUREC FEE
Degree Programme: BSc Hons in Computer Science
EcoCash Reference: MP240731.1348.L25270

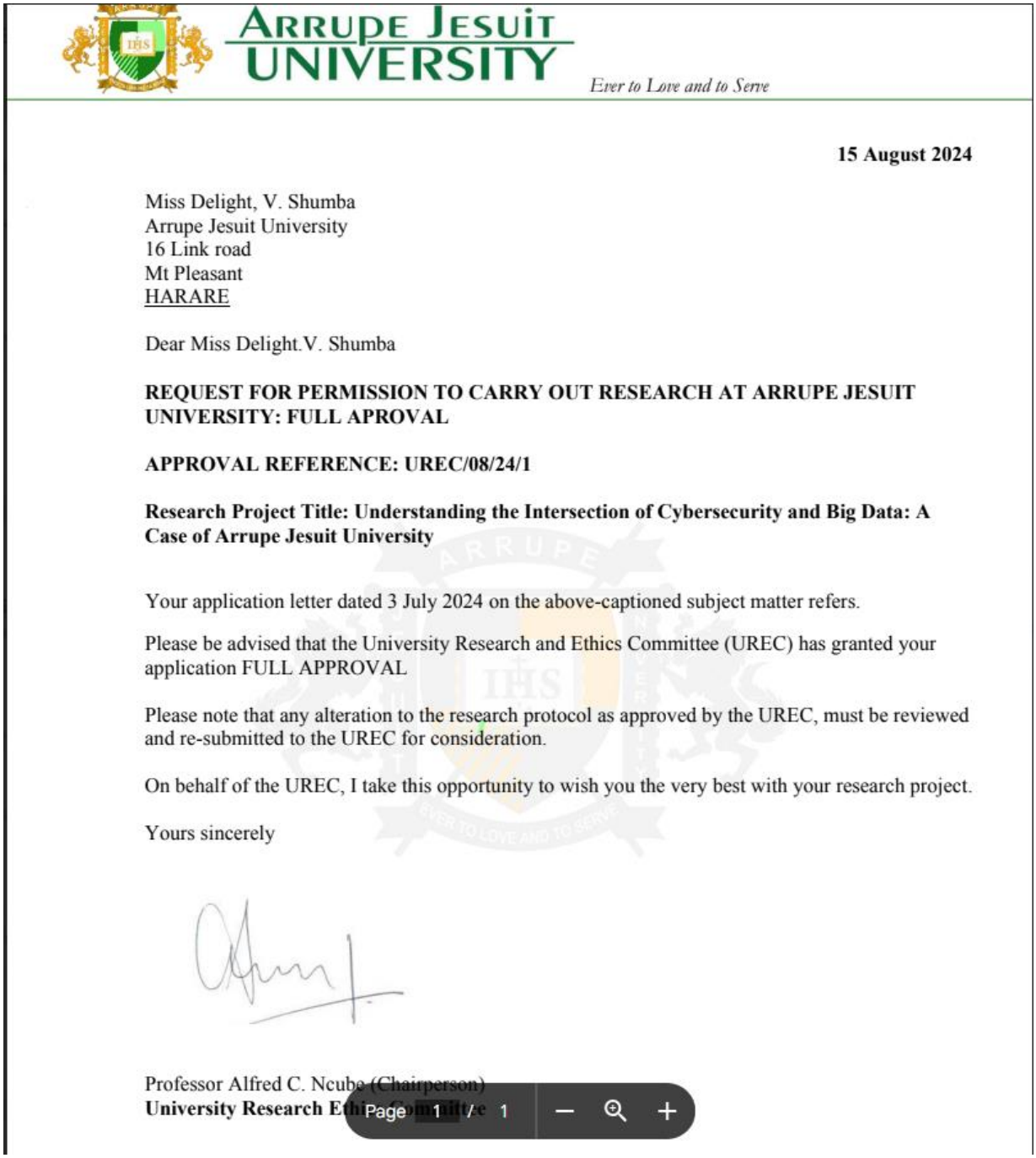
Your payment has been made to the payee.

To view the transaction online please follow this link

https://www.paynow.co.zw/Transaction/TransactionView/?guid=5ab20f89-311d-4c2f-b15c-45b51b79336b

Paynow is Zimbabwe's leading online payments platform. It's convenient and secure. Make and receive payments for

Ethical clearance letter



ARRUPE JESUIT UNIVERSITY

Ever to Love and to Serve

15 August 2024

Miss Delight, V. Shumba
Arrupe Jesuit University
16 Link road
Mt Pleasant
HARARE

Dear Miss Delight.V. Shumba

REQUEST FOR PERMISSION TO CARRY OUT RESEARCH AT ARRUPE JESUIT UNIVERSITY: FULL APROVAL

APPROVAL REFERENCE: UREC/08/24/1

Research Project Title: Understanding the Intersection of Cybersecurity and Big Data: A Case of Arrupe Jesuit University

Your application letter dated 3 July 2024 on the above-captioned subject matter refers.

Please be advised that the University Research and Ethics Committee (UREC) has granted your application FULL APPROVAL

Please note that any alteration to the research protocol as approved by the UREC, must be reviewed and re-submitted to the UREC for consideration.

On behalf of the UREC, I take this opportunity to wish you the very best with your research project.

Yours sincerely

Professor Alfred C. Ncube (Chairperson)
University Research Ethics Committee

Page 1 / 1

AUREC Approval letter

AFRICA UNIVERSITY
A United Methodist Related Institution

"Investing in Africa's future"

AFRICA UNIVERSITY RESEARCH ETHICS COMMITTEE (AUREC)

P.O. Box 1320 Mutare, Zimbabwe, Off Nyanga Road, Old Mutare-Tel (+263-20) 60075/60026/61611 Fax: (+263 20) 61785 Website: www.africau.edu

Ref: AU 3513/24 19 November 2024

DELIGHT V SHUMBA
C/O Africa University
Box 1320
MUTARE

RE: **UNDERSTANDING THE INTERSECTION OF CYBERSECURITY AND BIG DATA: A COMPREHENSIVE ANALYSIS AT ARRUPE JESUIT UNIVERSITY.**

Thank you for the above-titled proposal you submitted to the Africa University Research Ethics Committee for review. Please be advised that AUREC has reviewed and approved your application to conduct the above research.

The approval is based on the following.

a) Research proposal

- **APPROVAL NUMBER** AUREC 3513/24
  This number should be used on all correspondences, consent forms, and appropriate document
- **AUREC MEETING DATE** NA
- **APPROVAL DATE** November 19, 2024
- **EXPIRATION DATE** November 19, 2025
- **TYPE OF MEETING**: Expedited
  After the expiration date, this research may only continue upon renewal. A progress report on a standard AUREC form should be submitted a month before the expiration date for renewal purposes.
- **SERIOUS ADVERSE EVENTS** All serious problems concerning subject safety must be reported to AUREC within 3 working days on the standard AUREC form.
- **MODIFICATIONS** Prior AUREC approval is required before implementing any changes in the proposal (including changes in the consent documents)
- **TERMINATION OF STUDY** Upon termination of the study a report has to be submitted to AUREC.

AFRICA UNIVERSITY
RESEARCH ETHICS COMMITTEE (AUREC)
APPROVED
P.O. BOX 1320, MUTARE, ZIMBABWE

Yours Faithfully

MARY CHINZOU
ASSISTANT RESEARCH OFFICER: FOR CHAIRPERSON
AFRICA UNIVERSITY RESEARCH ETHICS COMMITTEE