

Data Security in Chatbots for the Insurance Industry: A case study of a South African Insurance Company

By

Zilungile Bokolo

Research Thesis submitted for review for the degree

Master of Information, Communication and Technology

In the Faculty of Informatics and Design

at the Cape Peninsula University of Technology

Supervisor: Prof Justine Olawande Daramola

Cape Town

Date submitted: June 2023

CPUT copyright information

The research proposal may not be published either in part (in scholarly, scientific or technical journals), or as a whole (as a monograph) unless permission has been obtained from the University.

DECLARATION

I, Zilungile Bokolo, declare that the contents of this dissertation/thesis represent my own unaided work and that the dissertation/thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.



01/06/2023

Signed

Date

ABSTRACT

As chatbots become more popular, the insurance industry has adopted their use. Although chatbot has been used a lot in customer relationship management (CRM), there is a lack of data security and privacy control strategies for data in chatbots. During data exchange, the client's data may be compromised through computer security breaches, thus exposing the client to possible fraud and theft. The lack of data security and privacy control strategies for data in chatbots has become a major security concern in financial services institutions. Chatbots access a lot of company and client information and that makes the data contained in chatbots to be the target of hackers which can cause harm to companies and customers.

This study explored how data security in chatbots in South African insurance organisations can be attained. To realise the aim of this study, five objectives were formulated as follows, to: 1) identify the potential use cases of chatbots for CRM in a South African insurance organisation; 2) identify the challenges of securing data in a chatbot in a South African insurance organization; 3) determine the security goals, threats, and vulnerabilities associated with the use of chatbots in a South African insurance organisation; 4) develop a threat model for the security and privacy of data in chatbots for a South African insurance organization; and 5) evaluate the threat model for security and privacy of data in the chatbots for a South African insurance organisation.

The mixed-methods research methodology was adopted for the study. A case study research strategy that involved data collection from a South African insurance company was used. Semi-structured interviews were conducted with participants that were purposively selected. Also, the STRIDE modelling approach was used to collect data on the security threats and vulnerabilities that pertain to each insurance use case with for each component of STRIDE — Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. Based on the outcome of the STRIDE modelling, a threat model for data security in chatbots for the South African insurance industry was developed using the Attack Defence tool. The threat model reveals the data security threats in chatbots, and how they can be mitigated. An evaluation of the threat model was conducted using security experts who assessed the quality of the threat model. They also provided qualitative feedback on the threat model. The evaluation of the threat model adopted the System Usability Scale (SUS) questionnaire which is a standard questionnaire to evaluate a system or product. The SUS score for each evaluator was calculated, and a mean SUS score was obtained.

From the expert evaluation, the developed threat model for data security in insurance chatbots obtained a mean SUS of 79.4 which corresponds to a grade B rating, which is a good rating based on the rules for the SUS scores. From the qualitative feedback, the security experts observed that the threat model can help to improve overall security and protect against potential attacks, and also proactively identify and mitigate potential threats in chatbots.

The insurance industry and academia will benefit from this study. Insurance organisations can implement security using the proposed threat model for the security of data in their business chatbots. Also, this study contributes new information to the body of knowledge since this is the first study to develop a threat model for data security in the chatbots in the context of the South African insurance industry using STRIDE modelling.

Key Words: Data Security, Artificial Intelligence, Chatbot, Insurance, Threats, Vulnerabilities, Threat Modelling.

ACKNOWLEDGEMENT

I wish to acknowledge:

First my God, for his love and for giving me the strength to carry out this research through to the end.

Second My supervisor Professor Justine Olawande Daramola for his patience, dedication, commitment, and mentorship throughout this research. Thank you so much Professor, for your contribution and the time you have invested in this research.

Dr Henri Knoesen for his continuous insightful advice, for providing help whenever I needed any, and for always being willing to share his knowledge.

Also, My manager, Mr Kedibone France Mangena, who has been supporting me throughout my postgraduate studies and always believing in me. I want to thank him for his guidance, motivation, and commitment.

And my mother, for her love, continuous support, prayers, and blessing throughout my life and academic career.

DEDICATION

I dedicate this thesis to the memory of my beloved brother, Azile Bokolo. Though he is no longer physically present, his spirit and influence continue to guide me in every step of my life. His belief in my abilities and constant encouragement propelled me forward, even in the face of challenges. I am forever grateful for the unwavering faith he had in me. He always emphasized the importance of education and the pursuit of excellence. His wisdom and guidance helped shape my values, shaping me into the person I am today.

PUBLICATION FROM THE THESIS

Journal Article

Zilungile Bokolo & Olawande Daramola (2023). Threat Elicitation for Data Security in Chatbot for the Insurance Industry. Informatics Journal (under review).

Paper in Preparation

Developing a Threat model for Data security in Chatbots using STRIDE: A Case Study of an Insurance Organisation

Table of Contents

DECLARATION i

ACKNOWLEDGEMENT iii

CHAPTER ONE 1

 INTRODUCTION..... 1

 1.1 Introduction..... 1

 1.2 Background..... 1

 1.3 Research Problem 2

 1.4 Aims, Objectives and Research Questions 3

 1.5 Delineation of Study 4

 1.6 Significance of the Study 4

 1.7 Thesis Structure 4

 1.8 Chapter Summary 4

CHAPTER TWO 5

 LITERATURE REVIEW 5

 2.1 Insurance Industry 5

 2.2 Chatbots Applications 5

 2.3 Chatbots Security 7

 2.4 Data Security 8

 2.5 Chatbot Architecture 8

 2.6 Threat Modelling 9

 2.7 Stride Modelling 10

 2.8 Related work 10

 2.9 Chapter Summary 15

CHAPTER THREE 16

 RESEARCH METHODOLOGY 16

 3.1 Research Approach 16

 3.2 Research Methodology 16

 3.4 Research Design 16

 3.5 Ethics of the study 19

 3.6 Chapter Summary 20

CHAPTER FOUR	21
DATA COLLECTION AND THREAT ELICITATION.....	21
4.1 Description of the Case Study	21
4.2 Data Collection	21
4.3 Findings from Analysis of Participant's Response	22
4.4 Dataflow Analysis on the Insurance Industry Chatbot	35
4.4.1 Signs of the Data Flow	35
4.5 Analysis of the Insurance Chatbot	36
4.5.1 iAssist	36
4.5.2 WhatsApp	42
4.6 Chapter Summary	44
CHAPTER FIVE	45
THREAT MODEL DEVELOPMENT	45
5.1 Modelling Security based on Stride.....	45
5.2 Notation Used for Representation of Insurance Industry Chatbots.....	46
5.3 Spoofing	46
5.3.1 Scenario Description of Spoofing	46
5.3.2 Threat Model for Spoofing	48
5.3.3 Description of the Threat Model for Spoofing	49
5.4 Tampering with Data	49
5.4.1 Scenario Description of Tampering with Data	49
5.4.2 Threat Model for Tampering with Data	51
5.4.3 Description of the Threat Model for Tampering with Data	52
5.5 Repudiation	53
5.5.1 Scenario Description of Repudiation	53
5.5.2 Threat Model for Repudiation	55
5.5.3 Description of the Threat Model for Repudiation	55
5.6 Information Disclosure	56
5.6.1 Scenario Description of Information Disclosure	56
5.6.2 Threat Model for Information Disclosure	57

5.6.3 Description of the Threat Model for Information Disclosure	58
5.7 Data Denial of Service	58
5.7.1 Scenario Description of Denial of Service	58
5.7.2 Threat Model for Denial of Service	60
5.7.3 Description of the Threat Model for Denial of Service	61
5.8 Elevation of Privilege	61
5.8.1 Scenario Description of Elevation of Privilege	61
5.8.2 Threat Model for Elevation of Privilege	64
5.8.3 Description of the Threat Model for Elevation of Privilege	65
5.9 Chapter Summary	66
CHAPTER SIX	67
THREAT MODEL EVALUATION	67
6.1 Profile of Evaluators	67
6.2 Evaluation Process	68
6.3 Evaluation Using System Usability Scale	68
6.4 Evaluation Results	69
6.5 Security Expert Feedback	70
6.6 Threat Validity.....	71
6.7 Summary	72
CHAPTER SEVEN	73
SUMMARY, CONCLUSION, RECOMMENDATIONS, AND CONTRIBUTION	73
7.1 Research Summary	73
7.2 Research Conclusion	75
7.3 Contribution to the Study	75
7.4 Future Research and Recommendations	76
REFERENCES.....	77
APPENDICES	82
APPENDIX A: Objective One and Two Interview Process	82
APPENDIX B: Threat Model Evaluation Questionnaires.....	83
APPENDIX C: Ethical Requirement.....	84

List of Figures

Figure 2.1: Electronic Frontier Foundation (EFF) Secure Messaging Score Card8

Figure 2.2: Architecture – Functional Bot9

Figure 2.3: Microsoft Threat Modelling Tool10

Figure 3.1: Research Design of the Study17

Figure 4.1: iAssist Use Case Diagram27

Figure 4.2: WhatsApp Diagram28

Figure 4.3: Critical Business Data Flow of iAssist Chatbot35

Figure 4.4: Login Process36

Figure 4.5: Claim Request37

Figure 4.6: Personal Lines Request38

Figure 4.7: Commercial Lines Request39

Figure 4.8: Human Resource Request40

Figure 4.9: WhatsApp Business Data Flow41

Figure 4.10: WhatsApp Bot User Interaction42

Figure 5.1: Spoofing Attack and Defence46

Figure 5.2: Tampering with Data Attack and Defence50

Figure 5.3: Repudiation Attack and Defence52

Figure 5.4: Information Disclosure Attack and Defence55

Figure 5.5: Denial of Service Attack and Defence58

Figure 5.6: Elevation of Privilege Attack and Defence63

List of Tables

Table 2.1: Types of Security Threats	11
Table 2.2: Related Work Reviewed	13
Table 3.1: Mapping of each objective to the specific stage.....	18
Table 4.1: Profile of Participants in the Study.....	20
Table 4.2: Findings from the Data Analysis.....	21
Table 4.3: Login Use Case Narrative.....	28
Table 4.4: Human Resource Use Case Narrative.....	29
Table 4.5: Claims iAssist Use Case Narrative.....	30
Table 4.6: Personal Lines iAssist Use Case Narrative	31
Table 4.7: Commercial Lines iAssist Use Case Narrative	32
Table 4.8: Policy WhatsApp iAssist Use Case Narrative	33
Table 4.9: Information Flow Signs	34
Table 5.1: Spoofing	44
Table 5.2: Tampering with Data	47
Table 5.3: Repudiation	51
Table 5.4: Information Disclosure	53
Table 5.5: Denial of Service	56
Table 5.6: Elevation of Privilege	59
Table 6.1: Thread Evaluation Participants	65
Table 6.2: Questionnaire for Threat Model Evaluation	66
Table 6.3: SUS Score Interpretation	68
Table 6.4: SUS Score	68

GLOSSARY

Terms/Acronyms/Abbreviations	Definition/Explanation
ADT	Attack Defence Tree
AI	Artificial Intelligence
CRM	Customer Relationship Management
CSF	Cyber Security Framework
DFD	Data Flow Diagram
EFF	Electronic Frontier Foundation
FAQ	Frequently Asked Questions
GDPR	General Data Protection Regulation
HR	Human Resource
ICT	Information Communication Technology
NIST	National Institute of Standards and Technology
OTP	One-time password
PL	Personal Lines
POPIA	Protection of Personal Information Act
RDS	Research Design Science
STRIDE	Spoofing, Tampering with Data, Repudiation, Information Disclosure, Denial of Service, Elevation of Privilege
SA	South Africa
SUS	System Usability Scale
UK	United Kingdom

CHAPTER ONE

INTRODUCTION

1.1 Motivation of the Study

The insurance industry provides a wide range of products that allows society to function in a safe environment by sharing the risk. These products include short-term (property and vehicle) and long-term (Life insurance) insurance; which would then cover both commercial and personal clients for their loss, at the claims stage. Insurers have been facing increasing competition from other industries such as banks, mutual funds, and investment advisory firms, which offer the same products (David, Sharon & Mary, 1998; Rebeena & Rosa, 2015; Ofori-Boateng, Ohemeng, Ahawaadong & Kwame, 2022). For the insurance companies, not to lose business chances to contenders, they needed to change how they were associated with their clients, by embracing the utilisation of chatbots. Previously, insurers used Customer Relationship Management (CRM) systems to manage their relationship with their customers. CRM frameworks normally centre around improving productivity and guaranteeing authority over the client by causing them to feel remembered for the business activities (Amber & Alan, 2017; Zhang , Li & Liu, 2023). Chatbots are increasingly adopted for various aspects of CRM in the insurance industry and it helps the company to minimize customer service costs, making the response process to the client very fast, and replying to up to 80% of routine inquiries (IBM, 2017).

However, there are issues concerning the security and privacy of data that are being transmitted between customers and Chabot, including customer information and the company's operation information. Chatbots have access to a lot of information that pertains to the company and clients which makes them the target of hackers, and other well-known web application security threats. Thus, there are security issues in chatbots (Josip & Franz, 2018, Tang & Nui, 2023).

This study aimed to investigate how to secure the information exchanged between clients and organizations via chatbots in the insurance industry. It is focused on developing a threat model to be used in ensuring the security and privacy of data in chatbots. The threat model of the study was developed from data collected through STRIDE modelling and the quality rating of the developed thread modelling was done by a security expert from the case study. The study is relevant because according to San (2019), companies need to improve their level of cybersecurity by successfully adopting a suitable cybersecurity strategy, which is necessary to obtain the level of cybersecurity required to protect the business, staff, clients, and reputation.

1.2 Background

The world of cyberspace has become a target of security breaches and information theft (NIST Framework, 2016). This issue causes trust concerns between organizations and clients. Nowadays, no

institution or organisation is without the use of Information Communication Technology (ICT) (Theodore, Andrew & Mirit, 2011; Samia, Muhammad & Umair, 2017). Most ICT applications require personal information and all the institution or business data including the personal data of users is generated electronically, with the help of ICT. Business transactions and the exchange of data between clients and organisations are performed electronically, and this presents a grave risk to users' private information, which can be easily stolen by hackers in compromised ICT security (Jashira, Noor, Siti & Nik, 2019). Technology advances in a way that all the information including clients, employees, and the organisation's operation becomes digitalised with the help of the internet but protecting or securing the information is still an issue (Sivaram & Abdullah, 2018). Marcus & Peter (2017) state that, with the accessibility of global systems and with the wide appropriation of PCs, the event of security strings and infringement turned into a mass phenomenon. Although the internet has opened a useful channel for communication it also introduced a new channel for criminals called cybercrime. Data breaches use computers and the internet to commit cybercrime. Sivaram & Abdullah (2018) state that criminal offences used to be committed before the computer age; now the same offences are being conducted using computers and smartphones. Jose, Wolfgang & Edward (2016) state that financial services institutions have always been vulnerable to fraud and are main targets for cybercriminals.

Currently, companies, including insurance and banks have taken the exciting next step in technology enhancement through the implementation of artificial intelligence (AI) within the business (Jahanzaib and Tarique, 2015). Ramnath, Ari & Doug (2018) in their study about the impact of AI on the future of insurance in 2030 state that "insurance will move from its present status of perceiving and repair to anticipate and prevent, transforming each part of the business in the process". The pace of progress will likewise quicken as brokers, financial representatives, insurers, etc. become more capable of utilising progressive innovations to improve dynamics and profitability, lower costs, and advance the client experience.

While there is a lot of excitement about Artificial Intelligence (AI) and other new technologies like chatbots, there seems to be significantly less understanding of securing these new technologies (Jahanzaib and Tarique, 2015). It is accepted that although the future of AI has a hugely useful power in the economy and everyday lives still it expands the issues identified with protection and security (Jahanzaib & Tarique, 2018). There is a need for chatbot security due to its AI features for the sake of customer information and personal information (Sen-Tarng, Fang-Yie & Jeng-Wei, 2018).

1.3 Research Problem

As chatbots become more popular the insurance industry has adopted their use (IBM, 2017; Oliver, 2019). Although chatbot has been used a lot in CRM; there is a lack of data security and privacy control strategies in Chatbots (Josip & Franz, 2018). During the data exchange, the client's data may be compromised through computer security breaches, thus exposing the client to possible fraud and theft.

The lack of data security and privacy control strategies in chatbots has become a major security concern in financial services institutions (Sen-Tarng, Fang-Yie & Jeng-Wei, 2018). Chatbots access a lot of

company and client information and that makes the information contained in chatbots to be the target of hackers and other well-known web application security threats which can cause harm to companies and customers (Josip et al., 2018).

The insurance organisation might be at a big loss if the data gets into the hands of hackers, which can have a very negative impact on the organisation (Jashira, Noor, Siti & Nik, 2019). Should important and sensitive information on insurance get lost, demolished, or falls into the hands of hackers it can lead to the loss of significant trade secrets or information that compromises confidentiality (Edin & Sejfudin, 2013).

Currently, there is no elaborate procedure to secure data in chatbots which makes insurance organisations, and customers vulnerable to security attacks.

1.4 Aim, Objectives, and Research Questions

This section describes the objectives that the study aimed to achieve and the research questions that guided the study to achieve its objectives.

1.4.1 Aim

The study aimed to explore how data security in chatbots in South African insurance organisations can be attained.

1.4.2 Objectives

The objectives of this research are described below:

1. To identify the potential use cases of chatbots for CRM in a South African insurance organisation.
2. To identify the challenges of securing data in a chatbot in a South African insurance organisation.
3. To determine the security goals, threats, and vulnerabilities associated with the use of chatbots in a South African insurance organisation.
4. To develop a threat model for the security and privacy of data in chatbots for a South African insurance organisation.
5. To evaluate the threat model for security and privacy of data in the chatbots for a South African insurance organisation.

1.4.3 Research Questions

To achieve the objectives of the study, the following research questions were formulated:

RQ1: How can data security in chatbots in the South African insurance industry be attained?

The sub-questions necessary to support the main research question are:

RSQ1: What are the potential use cases of chatbots for CRM in a South African insurance organisation?

RSQ2: What are the challenges of securing data in chatbots in a South African insurance organisation?

RSQ3: What are the security goals, threats, and vulnerabilities that are associated with the use of chatbots in a South African insurance organisation?

RSQ4: How can a threat model for the security and privacy of data in chatbots in South African insurance organisations be developed?

RSQ5: What is the quality rating of the threat model for security and privacy of data in the chatbot for South African insurance organisations from the perspective of relevant stakeholders?

1.5 Delineation of the Study

This study focuses on the insurance industry by using a case study of an insurance company in Cape Town, South Africa. Security concerns that pertain to other types of financial institutions apart from insurance are not covered.

1.6 Significance of the Study

The study is significant for looking at the bigger picture in terms of who/what is likely to benefit from the findings or product of the research and what the study will contribute (Ajaphol, 2007). This study will contribute to the improvement of data security in the insurance industry chatbots in South Africa. Organisations and academics will benefit from the study as the insurance organisations will use the data security threat model to implement security in chatbots and academics will benefit as this study contributes to the body of knowledge.

1.7 Thesis Structure

Chapter 1 contains the introduction and background and the research problem, objective, delineation, and significance of the study. Chapter 2 reviews the literature relevant to this study. It starts by discussing the chatbot applications, chatbot security, and the insurance industry in South Africa and then discusses thread modelling, which is described as a technique that is used in the identification of security threats. Chapter 3 presents the research methodology by discussing the research process used in this study. Chapter 4 presents the findings of the analysis of data collected from the case study. Chapter 5 presents the development of the proposed data security threat model. Chapter 6 presents the threat model evaluation process by security experts. Chapter 7 presents the summary, recommendation, and conclusion of the study.

1.8 Chapter Summary

This chapter introduced the research study. The research problem and the aim of the study are discussed. The research question and objective of the study were articulated in this chapter. The delineation and significance of the study are also discussed in this chapter.

CHAPTER TWO

LITERATURE REVIEW

This chapter presents a review of the literature that relates to this study. A literature review is a process of reviewing the previous work from scholars within the field of the proposed research and using that as a basis for data collection. It gives an outline of a subject field that bolsters the identification of particular research questions (Jennifer & Frances, 2004; Lim, Kumar & Ali, 2022).

2.1 Insurance Industry

The insurance industry is understood and known as a subdivision of financial services (Cummins and Doherty, 2006; Sibindi and Godi, 2014; Magano & de Beer, 2021). In South Africa, insurance companies are divided into two groups, namely long-term or life insurance and short-term property or car insurance (Sibindi and Godi, 2014; Magano et al., 2021). Good customer relationship management in the insurance industry is important as it keeps the existing customers, which can be done effectively through the adoption of advanced technologies (Roberts-Lombard, 2011; Ledro, Nosella & Vinelli, 2022).

Although customers can go directly to insurance companies, brokers or agents are often used as middlemen between insurance companies and customers. Brokers are also referred to as intermediaries; they always work closer to the clients to help them to understand the products offered by the insurance company. When the customer has decided on the product to take, the brokers then do a risk assessment and underwriting (Cummins and Doherty, 2006; Guillem, 2022). Therefore, intermediaries or brokers perform a huge role to give value to insurance clients, then they need support from the insurance industries by providing capabilities like a chatbot to improve the client's experience and retail more policies (Kanchinadam Qazi, Bockhorst, Morell, Meissner and Fung, 2019; Guillem, 2022).

Chatbots built on AI are an evolving technology (Riikinen et al., 2018; Lee, Bubeck & Petro, 2023) and the call centre employs these chatbots to support brokers. The implemented chatbot aims to improve call centre agent performance by answering questions posed by clients and that also helps the organisation to deliver a good service than its competitors (Meltzer, 2001; Zhang et al., 2023). Although there is excitement about the adoption of these new technologies in the insurance industry, security protection is a major concern, which makes insurance to be vulnerable to security breaches such as claim fraud, etc. (Mayank, Subhra, Sushmita, Sourav, Anupam & Kwok-Yan, 2019; Tang et al., 2023).

2.2 Chatbots Applications

A chatbot is a customer service platform, where clients get an opportunity to access information and help quickly and it also provides continuous customer service (Følstad, Nordheim & Bjørkli, 2018; Chien-Chang, Anna & d Stephen, 2023). Monthly, above 3 billion people use chatbots direct or

indirectly. Organisations have been putting resources into chatbot innovation to ensure competitive advantage by improving client support and diminishing expenses by 40% (Vagelis, 2018). In the case of the insurance sector, the execution of administration advancements dependent on chatbot innovation can contribute among different advantages to improve the effectiveness of the insurance value chain, diminishing costs, and producing client faithfulness and trust (Oliver, 2019; Zhang et al., 2023).

The talk about chatbots has been around since the 1950s. Alan Turing in the year 1950 started to challenge computers to find out if they can think. In 1966, Joseph Weizenbaum at MIT developed the first chatbot called ELIZA (Jack, 2017; Shobana, Kamireddy & Muthamsetty, 2023); however, as of late, there has been a fast increment in the number of chatbots due to some extent to a wide market selection of mobile and smart devices (Mengting, Paul, Perry & Vatche, 2016; Isinkaye, Imran & Michael, 2022). Grand view research (2017) in their reports states that the worldwide chatbot market will arrive at 1.23 billion US dollars in yield value in 2025. Chatbots are predicted to rescue about \$11 billion in companies with their increasing use (Juniper estimate , 2018; Sandy & Paula, 2022)

Nuruzzaman and Hussain (2018) discussed categories of chatbots from different studies. The first category discussed is from Chen, Liu, Yin, and Tang's (2017) study, where the category is described as task and non-task oriented. Task-oriented chatbots provide short interactions e.g. Apple Siri and Amazon Alexa, which are used for phone calls or travel directions, while a non-task-oriented chatbot is used for chat or conversation and contains questions and answers.

Nuruzzaman and Hussain (2018) then discussed the second category of chatbots from Barker's (2017) study, where the researcher categorised the chatbot into four; service chatbot, commercial chatbot, entertainment chatbot, and advisory chatbot. The service chatbots usually give support and service to the user. Commercial chatbots have the main purpose of giving support to the sales process or the communication of marketing messages. Entertainment chatbots are used for entertainment events and it is more informative. Advisory chatbots are implemented for recommendations, suggestions, advice, guidelines, etc.

Nuruzzaman and Hussain (2018) from their observations added another category to the two earlier mentioned by Chen, Liu, Yin, and Tang (2017) and Barker (2017) which they refer to as the third category of chatbots. It consists of four types: goals-based chatbots, knowledge-based chatbots, service-based chatbots, and response-generated-based chatbots. The goals-based chatbot executes specific tasks and short conversations, responds to the asked questions or tells the user if there is any problem. Knowledge-based chatbots are trained on specific data and depend on it to provide answers to the users. It usually contains public knowledge data. Service-based chatbots can satisfy personal or commercial necessities, such as providing commercial documents to the user as a replacement for a phone call to a call centre or sending emails. The fourth type of chatbot in this categorisation is a response-generated-based chatbot. This is the most complex type of chatbot and is classified on the action they perform when responding to the question. The input and output are natural language-based, and the responses are either template-based, generative, retrieved, or search-engine based.

2.3 Chatbots Security

Chatbots are mostly accessible through different platforms of messenger apps such as Facebook, and Skype and there is no proper security implementation on these platforms (Ondrisek, 2016; Bangera & Subrahmanya, 2023). Electronic Frontier Foundation (EFF) Secure Messaging Scorecard shows that Facebook Messenger and eight other messenger platforms are not secured in five of seven proven measurements. Figure 2.1 shows the security scores on different platforms of messenger apps. The topic was presented at the Privacy Week Conference in Vienna in a talk titled “Privacy and Data Security of Chatbots” and “Why you shouldn’t talk to your chatbot about everything”. Jack (2017), and Sakshi and Vinay (2023) states that although WhatsApp is the most secure messenger app and provides end-to-end encryption, should there be any failure on it, hackers can succeed in getting the data between users who are sharing the same network because they can perform sniffing and steal each other’s credentials. Messenger also does not confirm the identity of the user in an instance such as sending a One-time password (OTP). Previous chats are not hidden so if the hackers perform malicious attacks they can steal the credentials. Lastly, the security design is not well documented in these messenger apps.

Nathaniel (2018) states that between February and June 2018, there was a data breach that occurred in Ticketmaster’s global customer base which was discovered on 23 June. There was malicious software in their chatbot that was gathering information and sending it to a third party. The compromise of the chatbot made the customer’s confidential information including payment information to be stolen.

	Encrypted in transit?	Encrypted so the provider can't read it?	Can you verify contacts' identities?	Are past comms secure if your keys are stolen?	Is the code open to independent review?	Is security design properly documented?	Has there been any recent code audit?
Facebook chat	✓	✗	✗	✗	✗	✗	✓
Google Hangouts / Chat off the record	✓	✗	✗	✗	✗	✗	✓
iMessage	✓	✓	✗	✓	✗	✓	✓
Kik Messenger	✓	✗	✗	✗	✗	✗	✗
QQ	✓	✗	✗	✗	✗	✗	✓
Signal / RedPhone	✓	✓	✓	✓	✓	✓	✓
Skype	✓	✗	✗	✗	✗	✗	✗
Telegram	✓	✗	✗	✗	✓	✓	✓
Viber	✓	✗	✗	✗	✗	✗	✓
WhatsApp	✓	✓	✓	✓	✗	✓	✓

Figure 2.1: Electronic Frontier Foundation (EFF) Secure Messaging Scorecard (Ondrisek, 2016)

2.4 Data Security

These days, information size is expanding now and then from gigabytes to terabytes or even petabytes, basically as a result of the advancement of a lot of constant information. Big data is sent through the web and they are put away in the distributed computing environment. As distributed computing gives web-based administrations, there are numerous assailants and malicious clients. They generally attempt to get to clients' confidential huge information without having access rights. Sometimes, they supplant the original information with fake information. Therefore, data security has become a huge concern recently (Suyel, Debashree, Seifedine, Revathi & A, 2020). The leakage or alteration of data can be deliberate and unintended, and companies may be punished or held criminally liable for such incidents to ensure the privacy and integrity of the data is an active research area (Christian, Alfredo, Henry & Kim-Kwang, 2018; AbdulRaheem, Joseph, Chinmay, Emmanuel, Idowu & Akash, 2023).

2.5 Chatbot Architecture

Architecture is a fundamental structure that can be utilised for building up a wide range of applications (Khan, 2017; Poirier, Khalifa, & Wijidane, 2023). Roshan further states that because chatbots are rapidly adopted, organisations are forced to consider the related architecture approach when implementing the chatbot. Then categorise the chatbot architecture into three 1) A personal assistant, 2) A customer service bot, and 3) A functional bot. The study is more focused on functional bots since insurance is using the specialist bot, see Figure 2.2.

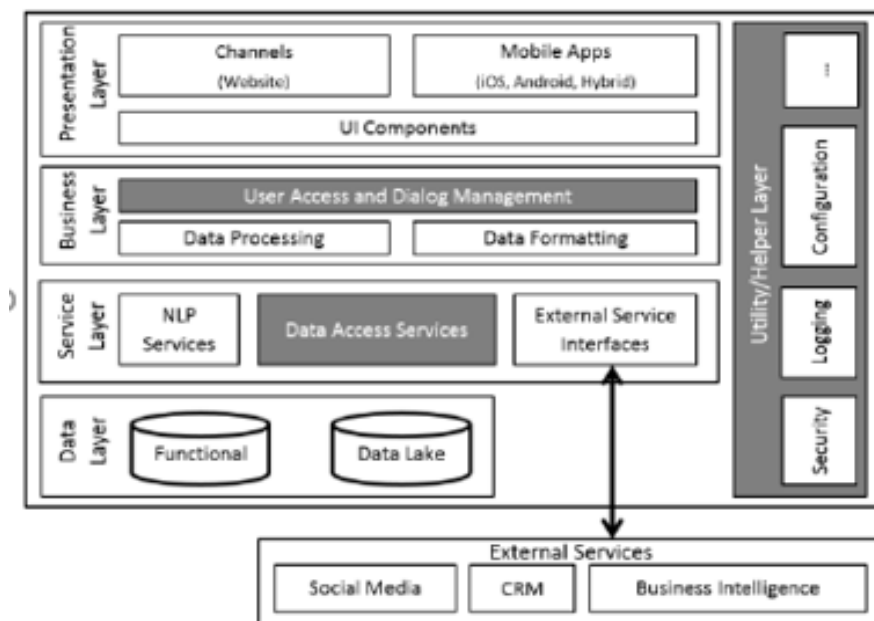


Figure 2.2: Architecture – Functional Bot (Khan, 2017)

In this architecture, the presentation layer contains the parts that actualise and show the user interface UI and oversee client association. In the business layer, data processing contains various stages that must be done in the right order. After processing, the data has to be structured according to the user

association. Service layer parts give access to both inner and outside information, business usefulness, middleware availability, and different administrations.

ChatBot solutions, because it is being presented to a huge number of frameworks, channels, and stages, itself makes it extremely vulnerable (Roshan, 2017; Edu, Mulligan, Pierazzi, Polakis, Tangil & Such, 2022).

2.6 Threat Modelling

The security of applications can be categorised into two; external and internal security. In a protected application, internal security is the main problem. This problem is determined by how security is implemented in the application design. This process includes identifying security threats in the application. Techniques like threat modelling have been implemented for the identification of security threats (Crothers, Japkowicz & Viktor, 2023). Threat modelling consists of several methodologies and techniques such as STRIDE (Lechner, Vjeran & Zlatko, 2023), Abuser stories (Crothers et al., 2023), STRIDE average model (Zaeni, Dyah, Anik & Muhammad, 2023), Attack trees (Ebrahimi, Christoph, Joaquim & Christoph, 2022), Fuzzy Logic (Batool, Mushtaq & Syed, 2022), SDL Threat Modeling tool (Santa, 2023), T-map (Hu, Ziqi, Yechao, Leo, Yifeng, Yuanyuan & Hai, 2022), and CORAS (Shafiq, Asif, Shabir, Ghulam & Sajid, 2014; Heisel & Marvin, 2023). Threat modelling is proposed as a goal for secure application improvement and framework security assessments. Its goal is to be progressively proactive and make it increasingly hard for aggressors to achieve malicious intent (Xiong & Robet, 2019; Crothers et al., 2023). According to Edin & Sejfudin (2013), the threat model for standardisation of data security ought to be lined up with the business strategy through successful usage, acquirement, and a combination of the framework. They also mention that through the use of ISO/IEC 27001, which is a worldwide set of principles for security, organisations can get useful guidelines. Microsoft defines threat modelling as a design method that can be used to assist with distinguishing threats, assaults, vulnerabilities, and countermeasures that could influence application. A view of Microsoft threat modelling is shown in Figure 2.3.

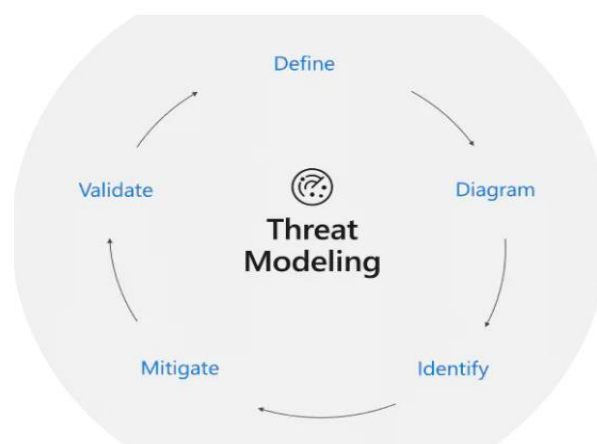


Figure 2.3: Microsoft Threat Modeling Tool (Microsoft, 2017)

2.7 STRIDE Modelling

The chatbot does not present new security concerns, since all the security concerns in the chatbot were previously discovered in other systems and appropriately mitigated. According to Paul (2019), security concerns in chatbots are categorised into two groups which are threats and vulnerability. Paul (2019) then used STRIDE Modelling to give different types of threats and attack intentions. STRIDE is a threat model methodology that is commonly used. STRIDE consists of six categories of threats which are Spoofing, Tampering, Repudiation, Information Disclose, Denial of Service, and Elevation of Privilege as shown in Table 2.1. In the STRIDE model, a data flow diagram of the application is implemented, and this model is used at each node of the Data Flow Diagram (DFD) of the application (Shafiq, Asif, Shabir, Ghulam & Sajid, 2014; Da Silva, Maxime, Pierre-Henri, Stephane & Nelson, 2023).

Table 2.1: Types of Security Threats (Paul, 2019)

No#	Category	Property	Attack Intention
1	Spoofing	Authentication	Illegal access and use of another user's credentials. Impersonating something or someone else
2	Tampering	Integrity	Aimed to maliciously change/modify data
3	Repudiation	Non-Repudiation	Aimed to perform an illegal operation in a system
4	Information Disclose	Confidentiality	Data theft
5	Denial of Service	Availability	Aimed to deny access to valid users
6	Elevation of Privilege	Authorisation	Aimed to gain privileged access

2.8 Related Work

Wube, Esubalew, Weldesellasiye & Debelee (2022) report indicates that chatbot security is still an issue, considering that user privacy, performance, and trust still appear to be the major factors impacting client satisfaction. The authors also indicate that the security and privacy vulnerabilities of chatbots in the financial sector must be considered and analysed before the developers deploy them. The limitation of the study is that the aim is broad as it does not only focus on security. The authors only indicated the security issues in the literature, with no suggestions or solutions in terms of mitigating the security threats or vulnerabilities in chatbots.

Hasal, Nowaková, Ahmed Saghair, Abdulla, Snášel & Ogiela (2021) state that chatbots are AI communication applications that are becoming progressively popular and most of the security questions in chatbots are not clearly solved. Chatbots are used for assistance in daily needs like online shopping, banking, and healthcare. However, it adds more security threats and produces critical security questions which need to be handled. Understanding the underlying issues requires identifying the critical steps in the methods used to design chatbots related to security. The authors talked about all the significant security, privacy, data protection, and social aspects of the usage of chatbots by reviewing the existing literature and producing a complete view of the given problem. Their study further indicates challenges in security and suggests ways to reduce security challenges found with the use of chatbots. The gap in the study is that although the authors presented the security issues concerning the use of chatbots, there is no proper existing framework used to identify those threats and mitigation steps. Also, there is limited information regarding the mitigation steps. STRIDE was just mentioned in the paper as a technique for the identification of threats and vulnerabilities but was not used.

Ng, Coopamootoo , Toreini , Aitken , Elliot & van Moorsel (2021) conducted a study where they investigated the effects of chatbot vignettes with socio-emotional features and without socio-emotional features with the intent to utilise the chatbot for financial support purposes. The study found that the social-emotional characteristics of chatbots in the financial industry can indicate a discrepancy between privacy and trust. The authors concluded that a suitable precautionary analysis concerning the security and privacy vulnerabilities of a chatbot in the financial industries must be executed before deployment. The researchers approached the study by using a Vignette-style methodology as an induction protocol. The gap or limitation of the study is that the study did not utilise actual chatbot prototypes and socio-emotional cues integrated with financial systems.

Saiful, Abdur, Sadek, Mohammad, Mohammad & Sasu (2020) integrated chatbots and blockchain technology and the reason was to improve chatbot security issues in the financial sector. The authors implemented a proof of concept and did the evaluation of performance and analyses of several security and privacy concerns conducted by applying a blockchain-enabled chatbot. The researchers approached the study by formulating a list of requirements based on rigorous threat models for chatbots in the financial sector. The authors provided a proof-of-concept prototype and defined its protocol flow to demonstrate applicability.

Ye and Li (2020) examined the security and privacy vulnerabilities of existing chatbots and proposed that to avoid substantial harm, chatbot developers should perform a security analysis before any deployment. The authors analysed potential security and privacy exposures in the chatbot architecture and discovered that the chatbot community has not yet implemented comprehensive requirements for chatbot security. The authors approached the study by first understanding how the existing chatbot architecture works. They achieved that by following the path that a message takes from the client module to the communication module, to the response generation module, and to the database module and then identifying possible attacks in each module. There is no framework used in identifying chatbot security attacks and no mitigations. The study suggested that future work can pay attention to discovering attacks that span from one module to a module of the chatbot architecture.

Lai, Leu & Lin (2018) state that chatbots with artificial intelligence features may encroach on client security and individual protection. Security has become a significant issue that chatbots must focus on. Their study aimed to develop the Chatbot Security Control Procedure (CSCP) for banks to monitor the security of chatbots and ensure the protection of clients. The findings of their study show that there is no security in the chatbot and the security loophole in chatbots is caused by the AI security software.

In 2016, Microsoft CEO Satya Nadella proposed that the security improvement of artificial intelligence must entail six values of which two values must be added to a chatbot plan or strategy. The two values are i) Artificial Intelligence needs to be transparent and ii) Artificial Intelligence needs intelligent privacy.

Følstad, Nordheim & Bjørkli (2018) state that security and privacy in a chatbot are something that we must pay attention to. Their study investigated the initial set of issues assumed to be factors affecting clients' trust in chatbots for client service. The findings from the study show that the main issue of the clients, not trusting chatbots is because of their poor security and privacy. The researcher used an exploratory research design and then conducted a semi-structured interview.

Harkous, Shin, Fawaz & Aberer (2016) developed a PreBot that allows privacy within a conversation or chat between the user and chatbot. The reason to develop the privacy conventional bot is that they had a concern that the current chatbots are failing to protect users' privacy. The PreBot has an interface that provides the user with privacy settings and a set of privacy policies for the service provider.

Magdalene, Kovila, Ehsan, Mhairi, Karen & Aad (2020) state that there is still a user's private information concern and risk in a financial bot especially when it comes to the payment process because users have to disclose sensitive information like credit card information. These concerns can lead to users not trusting and using financial chatbots at all. The study focused on trust, privacy concerns, and social presence in using the chatbot. It was a quantitative study that was conducted in the United Kingdom (UK).

Krishna, Sergey & Pavol (2020) state that although there are more than 2000 chatbots in the market service, vendors do not prioritise the security risk of chatbots. The study made an example of Ticketmaster's global customer base chatbot which was affected by a data breach in 2017 and Delta Airlines' chatbot, in which hackers succeeded in stealing customer payment data. The study specified the important risks associated with chatbots like information leaks, denial of service, wrong advertisement, spam, malware, phishing, DDoS, man-in-the-middle attacks, data gathering, etc. which are necessary to be addressed. The study proposed Service Level Agreements as a solution to manage the risk regarding the chatbot service.

Summarily, several researchers have expressed their concerns about the security risk in chatbots, some researchers' efforts focused on chatbot security threats and vulnerability investigation without presenting how those threats must be mitigated, and some have looked at both chatbot security threats and mitigation. However, none of the previous studies has focused on data security in South African insurance chatbots, and none of the studies used STRIDE modelling in identifying security threats in chatbots and developed a threat model for insurance chatbots. This study is different from previous

studies as it introduced STRIDE as a proper way of identifying security threats in insurance chatbots and developing a threat model for chatbots that are used in the insurance industry.

Thus, the current study contributes to the new body of knowledge as it used STRIDE modelling to identify all the threats, vulnerabilities, and mitigations in insurance chatbots, and developed a threat model based on the information collected from STRIDE modelling that can enable the security of data in chatbots used in the insurance industry.

Table 2.2: Summary of Related Studies

Author	Aim/Objective	Findings	Gap/Critique
Wube, Esubalew, Weldesellasie & Debelee (2022)	Major factors impacting client satisfaction	chatbot security is still an issue, considering that user privacy, performance, and trust still appear to be the major factors impacting client satisfaction	the aim is broad as it does not only focus on security, but the authors also only indicated the security issues in the literature, with no suggestions or solutions in terms of mitigating the security threats or vulnerabilities in chatbots.
Hasal, Nowaková, Ahmed Saghair, Abdulla, Snášel & Ogiela (2021)	Understanding the underlying issues requires identifying the critical steps in the methods used to design chatbots related to security	Reported the challenges in chatbot security and suggests ways to reduce security challenges found with the use of chatbots.	Although the authors presented the security issues concerning the use of chatbots, there is no proper existing framework used to identify those threats and mitigation steps. Also, there is limited information regarding the mitigation steps. STRIDE was just mentioned in the paper as a technique for identifying threats and vulnerabilities but was not used.
Ye and Li (2020)	Analysis of potential security and privacy exposures in the chatbot architecture	discovered that the chatbot community has not yet implemented comprehensive requirements for chatbot security.	There is no framework used in identifying chatbot security attacks and no mitigations. The study suggested that future work can pay attention to discovering attacks that span from one module

			to a module of the chatbot architecture.
Ng, Coopamootoo , Toreini , Aitken , Elliot & van Moorsel (2021)	investigated the effects of chatbot vignettes with socio-emotional features and without socio-emotional features with the intent to utilise the chatbot for financial support purposes.	The study found that the social-emotional characteristics of chatbots in the financial industry can indicate a discrepancy between privacy and trust. The authors concluded that a suitable precautionary analysis concerning the security and privacy vulnerabilities of a chatbot in the financial industries must be executed before deployment.	The gap or limitation of the study is that the study did not utilise actual chatbot prototypes and socio-emotional cues integrated with financial systems.
Ng, Coopamootoo , Toreini , Aitken , Elliot & van Moorsel (2020)	Trust, privacy concerns, and social presence in using the chatbot.	Concern and risk in a financial bot more especially when it comes to the payment.	This study does not come up on what should be done to mitigate the security risk in chatbots.
Følstad, Nordheim & Bjørkli (2018)	Investigate and recognise an initial set of issues assumed to be factors affecting clients' trust in chatbots for client assistance.	Security concerns. For future research security and privacy, it's something that must pay attention to.	Critical knowledge gap on what can be done to prevent security risks.
Lai, Leu & Lin (2018)	Develop the Chatbot Security Control Procedure (CSCP) to monitor the security of the chatbot and ensure the protection of clients.	No security in the chatbot. A security loophole in chatbots brought by AI security software.	No chatbot security guidelines for insurance because this study was only based in the bank sector.

Harkous, Shin, Fawaz & Aberer (2016)	Developed a PreBot that allows privacy within a conversation or chat between the user and the chatbot.	Current chatbots are failing to protect users' privacy.	PreBot has an interface that provides the user with privacy settings and a set of privacy policies for the service provider. No security guidelines were followed in implementation.
--------------------------------------	--	---	--

2.9 Chapter Summary

This chapter reviewed literature relevant to this study. The types of the insurance industry in South Africa were defined. The types of chatbots and relevant architecture were defined. The concerns for data security in chatbots were explained. Methods to properly identify and mitigate security vulnerabilities and threats were clearly defined. Literature on what is already done in this area of research was reviewed. Literature on existing methods regarding data security in chatbots was defined.

CHAPTER THREE RESEARCH METHODOLOGY

This chapter presents the description of the methodology adopted for this study. Research methodology directs the researcher in determining what type of information is necessary for a study and which data collection tools will be most suitable for the study (Adil & Khalid, 2016). According to Ranjit (2016),

research methodology provides researchers with methods to determine answers to research questions. This is what this chapter presents.

3.1 Research Approach

The inductive approach is theory-building research that forms a theory from qualitative empirical data. It begins with the observations, and hypotheses are proposed when the research process is about to finish as the outcome of observations (Wayne & Stuart, 2004; Varpio, Elise & Sebastian, 2020). The inductive approach's goal is to produce meanings from the informational index gathered so that a researcher can classify patterns and associations to form a theory. In this study, there is a pursuit of an in-depth understanding of security challenges that pertain to the use of chatbots in the insurance industry and thereafter a proposal of a threat model to address the identified potential security challenges. Thus, this study employed the inductive approach.

3.2 Research Methodology

Research can be broadly categorised into quantitative, qualitative, and mixed methods methodologies; the methodology used for a particular study is strongly determined by the research question. In qualitative studies, results are called findings. The inductive approach enables the researcher to explore and understand a phenomenon better. The data in quantitative research is presented in numerical form or data can be converted into statistics. Mixed methods methodology is a combination of both qualitative and quantitative methods (Christopher & Rechar, 2013; Taherdoost, 2022). This study employs a mixed methods methodology since the study dealt with the collection and analysis of both qualitative data and quantitative data. The interviews and document review, and threat elicitation were qualitative data, while the rating of the threat model using SUS was quantitative data. Also, the evaluators gave feedback in the form of qualitative comments at the end hence the study adopted the mixed methods methodology.

3.3 Research Design

A research design is a logical order which links the empirical data to a study's research questions and, finally to its conclusions (Yin, 2018; Muzari, Goerge, & Samantha, 2022). Research design (RD) is defined as a tool to assist researchers to examine the research questions of a study. Types of qualitative research designs are ethnography, phenomenological, and case study. There are two types of case studies, which are single-case studies, and multiple-case studies. In multiple case studies, the researcher is studying more than one case to comprehend the distinctions and similarities between the cases. In a single case study, the researcher is interested in studying one single thing (Johanna, 2017; Muzari et al., 2022). In this study, a single-case study is implemented because the study is based on data security in the chatbot at a South African Insurance organisation, as shown in Figure 3.1.

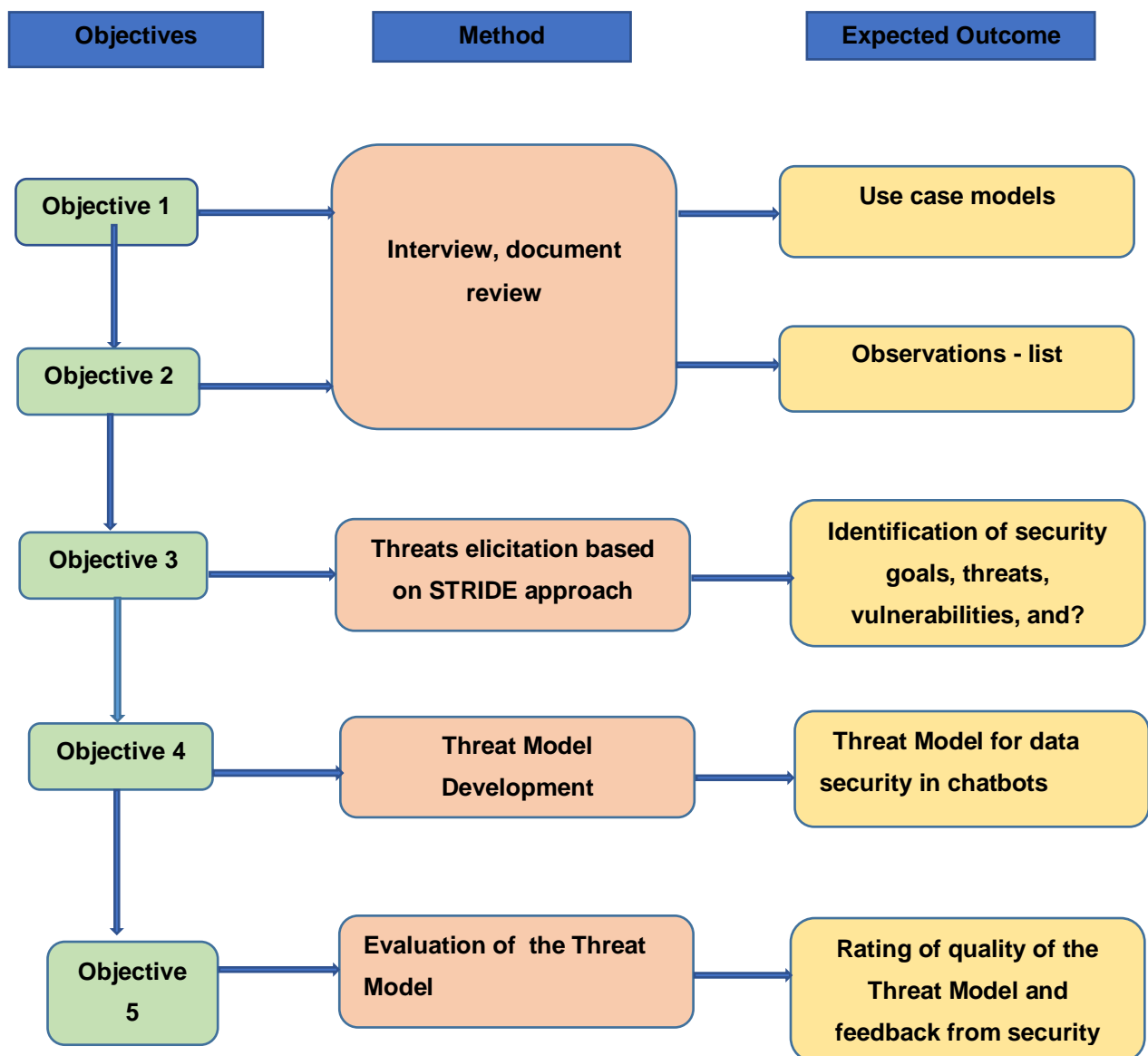


Figure 3.1: Research Design of the Study

The design of the research was organised in the following manner:

Stage 1 of RD: Interview, document review

At this stage, the researcher conducted semi-structured interviews in the process of collecting the data. This was done through the case study of a South African insurance company. The participants in this study were purposively selected as follows: security experts, chatbot developers, testers, and chatbot users. The literature related to the study was reviewed to guide and support the researcher's views. The expected outcome at this stage was use-case models.

Stage 2 of RD: Interview, document review

At this stage, the researcher conducted semi-structured interviews in the process of collecting the data. This was done through the case study of a South African insurance company. The participants in this study were purposively selected as follows: Security experts, chatbot developers, and testers, chatbot users. The related literature to the study was reviewed to guide and support the researcher's views

Stage 3 of RD: Threat Elicitation

At this stage, STRIDE modelling was used as an approach for threat elicitation, which is a dispensable step for the identification of threats and vulnerabilities associated with data security in chatbots. The results of the study show that all the chatbot use cases are vulnerable to each component of STRIDE modelling. The participants for data collection at this stage were five security experts from the case study

Stage 4 of RD: Threat Model Development

At this stage, a threat model for data security in a chatbot for a South African insurance organisation was developed. This was done through the data collected with the STRIDE model. The threat model was developed from the Attack Defence Tool called ADTool which is used to show the actions of an attacker trying to compromise the system and possible counteractions of a defender trying to protect the system, it also helps with the qualitative analysis of security, using attack–defence trees

Stage 5 of RD: Evaluation of the Threat Model

At this stage, the security experts evaluated the threat model using System Usability Scale (SUS) which is a tool to evaluate the system or product. Throughout the evaluation process, the questionnaire was designed using the SUS approach to assess the threat model proposed for data security in chatbots within the insurance industry based on STRIDE. After the evaluation process, the study shows evaluation results which is a SUS score. The security experts also gave general comments and observations (in the form of a narrative) on the quality of the threat model.

The demonstration of the mapping of each objective to the specific stage of research design is shown in Table 2.

Table 3.1: Mapping of each objective to the specific stage

Stage	Objective	Method	Expected output
1	To identify the potential use cases of chatbots for CRM in	Interview, document review	Use case models

	the South African insurance industry.		
2	To identify the challenges of securing data in chatbots in the South African insurance industry.	Interview, document review	Observations – list
3	To determine the security goals, threats, and vulnerabilities associated with the use of chatbots in the South African insurance industry.	Threat modelling using a STRIDE modelling approach	Identification of goals, threats, vulnerabilities, and mitigation strategies
4	To develop a guideline for data security and data privacy in chatbots for the South African insurance industry.	Data analysis; Guideline formulation	Recommendations, models
5	To evaluate the proposed guideline for the security and privacy of data in the chatbot for the South African insurance industry.	Evaluation of the threat model by security experts	Rating of quality of the Threat Model and feedback from security experts

3.4 Ethical Considerations

Ethics in research generally means a researcher has to protect everyone participating in the study from any loss or harm, protect participants' private information or privacy, etc. (Anon, 2007; Alwahaby, Cukurova, Papamitsiou & Giannakos, 2022). The below subtopics discuss how the ethical risks associated with this research will be mitigated:

Informed Consent: Informed consent is a course in which a researcher informs the participant about the nature, actions, risks benefits, etc. of the research in a way that is not technical and for easy understanding by the participants in the research (Sil & Das, 2018; Alwahaby et al., 2022). Everyone participating in this research received a consent letter which is a written communication from the Cape Peninsula University of Technology that explains that the researcher has the approval of collecting information to carry out the study.

Confidentiality of Participant's Information: The researcher is collecting private and personal information; therefore, the researcher must preserve privacy and restrict illegal access to the study data (Wolf et al., 2015; Alwahaby et al., 2022). The research data included signed consent letters, transcripts, and notes that were taken during the interview. personal logs are kept on a personal laptop and they will be kept for five years.

Security of Data for Companies and Individuals: The benefit of using unidentified identifiers can assist in protecting the identity of each participant or organisation participating in the research (Barocas & Nissenbaum 2014; Alwahaby et al., 2022). The study used anonymous coding (Participant 1, Participant 2, Participant 3 ..., etc.) to link each participant to their information. Password is used to save the data that was collected from participants to avoid illegal access.

3.5 Chapter Summary

This chapter explained the research methodology adopted for this study. The research philosophy, design, and approach that guided the research process were explained; data collection and analysis techniques that were used when collecting and analysing data for this study were also explained.

CHAPTER FOUR

DATA COLLECTION AND THREAT ELICITATION

In this chapter, the process of threat elicitation, which involves collecting relevant data from the organisation to identify security goals, threats, vulnerabilities, and mitigation strategies is reported. The first round of interviews was conducted to get the insurance chatbot use cases. The second interview was conducted through STRIDE modelling to identify how each use case is vulnerable to each component of STRIDE modelling.

4.1 Description of Case

The case study of the research is the insurance industry in South Africa. It is one of the leading insurance industries in South Africa. The size of the organisation is medium to large. It provides short-term insurance for properties, businesses, and cars. Its clientele is businesses and individuals.

4.2 Data Collection

A semi-structured interview is the primary data collection method used in this research to develop the proposed threat model for data security in chatbots. Semi-structured interviews are a data collection method commonly used in qualitative research. Researchers have a list of themes and possible key questions to be covered (Louise & Alison, 1994; Kallio, Pietila, Johnso & Kangasniemi, 2016; Naz, Gulab & Aslam, 2022). The study used semi-structured interviews because the researcher interviewed participants that are knowledgeable about insurance and how chatbots are used in the insurance industry. The roles and officers that were purposively selected to participate in the research are as follows: security expert, chatbot developers, chatbot testers, chatbot users, and chatbot managers. The study also collected data using STRIDE modelling, document review and also reviewed the literature. STRIDE modelling is a threat elicitation to identify security goals, possible threats, and vulnerabilities of each chatbot in the organisation regarding the different components of STRIDE based on the perspectives of the security expert. The experts were presented with a document template to capture their individual analysis of security goals, threats, and vulnerabilities, and thereafter areas of consensus were noted, while areas of differences were resolved in joint meetings of the experts. This led to the final documentation of the security goals, threats, and vulnerabilities associated with the use of chatbots in the insurance organisation.

S/N	Role	Number	Reference
1	Security Experts	2	A, D

2	Chatbot Developers	2	B, E
3	Chatbot Testers	2	F, G
4	Chatbot Users	2	C, J
5	Chatbot Managers	2	H, i

**Table 4.1:
Profile of
Participants in
the Study**

4.3 Findings from the Analysis of Participant's Responses

Table 4.2 shows the analysis of the data collected from the case study. Objectives one and two of the study were achieved in this section. The goal column represents the grouping of themes of the data. Alphabets A to J represent participants of the study during data collection.

Table 4.2: Findings from the Data Analysis

S/N	Goal	Finding from Participant's Responses
Objective 1: To identify the potential use cases of chatbots for CRM in a South African insurance organisation.		
1	The purposes of a chatbot in the organisation	<p>iAssist:</p> <ul style="list-style-type: none"> - Manage a relationship with clients effectively - Provide a faster and easy way of finding information - Improve productivity - Reduce training to contact centre agent - Less expense - Consistency of response - Easy to use - Allows information search <p>Participant (A, C, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Manage a relationship with clients effectively - Cut out the middleman manual intervention of customers having to speak directly with consultants - Provide self-service

		<ul style="list-style-type: none"> - Easy and fast access to information - Consistency of response - Easy to use - Provide 24/7 access Participant (A, B,D, F, G, & J)
2	Purpose fulfilled by the chatbot in the organisation and how the organisation benefits positively from using the chatbot.	iAssist: <ul style="list-style-type: none"> - Gives answers to the employees on repetitive questions. - Guide contact centre agent on how to ensure certain assets - Easy and quick access to information - Improve productivity - Give support to contact centre agents in helping clients - Consistency of response - Easy to use - Improves contact centre agent performance Participant: (A, B, C, E, F, G, H, I & J) WhatsApp bot: <ul style="list-style-type: none"> - Manage a relationship with clients effectively - Cut the middleman intervention - Easy and fast access to information - Consistency of response - Easy to use - Provide 24/7 access Participant: (A, B,D, F, G, & J)
3	The specific processes or operations where chatbots are used in the organisation	iAssist: <ul style="list-style-type: none"> - It gives the following functionalities: Personal Lines, Commercial Lines, Claims, and Human Resource - Staff ask HR-related questions like annual leave inquiries, Pension Fund, etc - HR bot also contains policy documents like the Organisation's Security Policy, Social Media Security Policy, etc. - Personal Lines and Commercial Lines bot provide policies offered, underwriting guidelines and rules, details of allowed risk, and items that are allowed to be insured Participant: (A,D, F, G, & J) WhatsApp bot: <ul style="list-style-type: none"> - Provide the following functionalities: Send my policy schedule, Send my confirmation of cover, border letter Participant: (A, B,D, F, G, & J)

4	The kind of support and maintenance available for chatbots in the organisation	<p>iAssist:</p> <ul style="list-style-type: none"> - Has learning capability - Keep all the asked questions and learn from them - Build on chatbots as information changes and as it starts moving into other areas of the business - Building upon putting in new content - There is a team that always checks how many questions were asked, and how many were answered or not answered by the chatbot. - The service provider provides first-line support, changes existing frequently asked questions and answers, and adds or deletes, or updates documents - The service provider provides business and technical support as well as maintenance for the solution, and there are various types of maintenance, preventative maintenance, adaptive maintenance, etc. <p>Participant: (A,C, D, E, F, G, H , I& J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Keep and route all the unanswered questions to the available agent - If the agent does not know the query it takes it up with the technical team <p>Participant: (B & E)</p>
5	Kind of data stored in chatbots in the organisation	<p>iAssist:</p> <ul style="list-style-type: none"> - HR-related policy documents - HR-related information - Underwriting rules and information - Details of allowed risk - Items that are allowed to be insured - Client and employee information <p>Participant: (A,C, D, E, F, G, H , I& J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Policy documents - Policy-related information - Client information <p>Participant: (A, B,D, F, G, & J)</p>
6	Chatbot users	<p>iAssist:</p> <ul style="list-style-type: none"> - Employees - Call centre agent <p>Participant: (A,B, C, D, E, F, G, H, I & J)</p>

		WhatsApp bot: <ul style="list-style-type: none"> - Clients - Call centre agent Participant: (A,B, C, D, E, F, G, H, I & J)
Objective 2: To identify the challenges of securing data in a chatbot in a South African insurance organisation.		
1	A place where the chatbots used in your organisation are hosted	iAssist: <ul style="list-style-type: none"> - Third-party cloud storage Participant: (A, , F, G, & J) WhatsApp bot: <ul style="list-style-type: none"> - Client's Third-party cloud storage - Structured database Participant: (A, B & E)
2	The integration of the organisation's chatbot into social media	iAssist: <ul style="list-style-type: none"> - It's a web bot, not integrated into the social media platform Participant: (A, D, F, G, & J) WhatsApp bot: <ul style="list-style-type: none"> - It's integrated via WhatsApp Participant: (A, B & E)
3	The data storage area for the chatbot platform in the organisation after transactions	iAssist: <ul style="list-style-type: none"> - Are stored in a third-party Mongo database Participant: (A, D, F, G, & J) WhatsApp bot: <ul style="list-style-type: none"> - Are stored in a third-party Mongo database Participant: (A, B & E)
4	Used data inside chatbots	iAssist: <ul style="list-style-type: none"> - Keep for reviews, optimisation purposes, and ongoing maintenance of the solution Participant: (A, D, F, G, & J) WhatsApp bot: <ul style="list-style-type: none"> - Keep for reviews, optimisation purposes, and ongoing maintenance of the solution Participant: (A, B & E)
5	Features of the Chatbot	iAssist: <ul style="list-style-type: none"> - Text-based Participant: (A, B, C, D, E, F, G, H, I & J) WhatsApp bot: <ul style="list-style-type: none"> - Text-based Participant: (A, B, C, D, E, F, G, H, I & J)

6	The security measures chatbots have to prevent identity theft	<p>iAssist:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to Protection of Personal Information Act and any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to POPOA; any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p>
7	The kind of security the chatbots have to ensure data privacy	<p>iAssist:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to POPOA; any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to POPOA; any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p>
8	Kind of security measures chatbots have to ensure data integrity	<p>iAssist:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to POPOA; any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption - The security measure in the chatbot to ensure data integrity comes under the same controls that are implemented for POPIA and With the GDPR in terms of

		<p>making sure the person is who they say are</p> <ul style="list-style-type: none"> - Encryption of information at rest <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Chatbots data is subject to privacy legislation so parts that are hosted here in South Africa are subject to POPOA; any other part that is hosted in Europe is subject to GDPR from a data privacy point of view - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p>
9	Kind of security measures chatbots have to prevent unauthorised access	<p>iAssist:</p> <ul style="list-style-type: none"> - Users can only access it once authenticated herself - A user signing in with your staff credentials - End-to-end encryption - <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - Users can only access it once authenticated herself - It validates that at least the details the user provides are those details on the policy that the organisation has. - End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p>
10	Kind of security measure chatbots have for user authentication	<p>iAssist:</p> <ul style="list-style-type: none"> - The staff signs in with the username and password, which is the security measure for user authentication - The authentication is provided by roles that are stored and implemented in the application and it has 3 roles menu with end-user content, administrator, or manager and those roles assist in authorisation in terms of which department the user got access to, as well as which reports are accessed End-to-end encryption <p>Participant: (A, B, C, D, E, F, G, H, I & J)</p> <p>WhatsApp bot:</p> <ul style="list-style-type: none"> - validate based on something that the person knows and therefore use that to prove that the user is authentic. - The user adds an ID number as a user identification method and the chatbot sends the OTP to the user. After the user

		<p>has confirmed, it sends the information requested to a user's email.</p> <ul style="list-style-type: none"> - If the user enters the OTP number incorrectly, it will just take the user back to the main menu. If the user does not have active policies linked to the ID number, then the bot will tell the user that no active policies are linked to this ID number. - Also, when the chatbot sends back personal details like phone number, ID number, etc. it masks e.g. 083*****92, mil****ab@gmail.com - End-to-end encryption <p>Participant (A, B, C, D, E, F, G, H, I & J)</p>
11	The security vulnerabilities that have been found with chatbots	No variabilities were found

Use case diagrams in Figures 4.1 and 4.2 show the use case of chatbots in the insurance company. iAssist has five use cases as follows: Personal Lines, Commercial Lines, Claim, Human Resource, and Login. It is being accessed by contact centre agents, administrators, and employees. WhatsApp has five use cases as follows: Policy Schedule, Confirmation of Cover, Border Letter, Home Agent, and User Authentication. WhatsApp is being accessed by contact centre agents, administrators, and clients. Tables 4.2 to 4.7 show chatbot use case narratives of chatbot use cases.

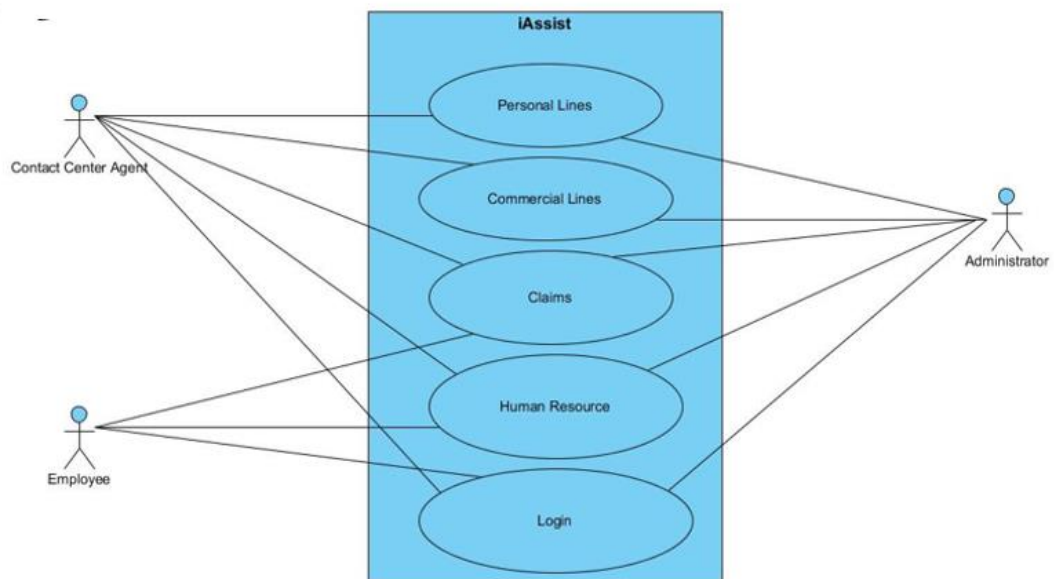


Figure 4.1: iAssist Use Case Diagram

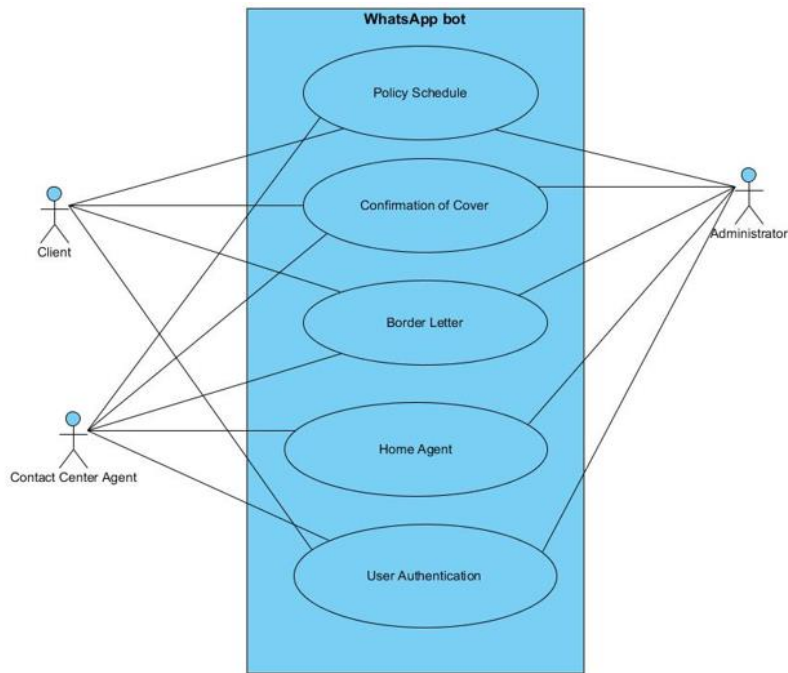


Figure 4.2: WhatsApp Use Case Diagram

Table 4.3: Login Use Case Narrative

USE CASE NAME: LOGIN	
PRIMARY ACTORS	Employee, Contact Centre Agent, Admin
SECONDARY ACTORS	
DESCRIPTION	User Log-in into the iAssist Chatbot
USER ACTION	IASSIST RESPONSE
1. User Enters Username and Password	2. Validates if textboxes are not empty
	3. Checks if the username and password match
	4. Login the user into the Chatbot
ALTERNATE	
Step 2 Textboxes are empty	Notifies the User to fill in mandatory fields
Step 3 Username and Password don't match	Alerts the User that the Username or Password is incorrect. Prompt the User to enter the username and password again.

Table 4.4: Human Resource iAssist Use Case Narrative

USE CASE NAME: HUMAN RESOURCE IASSIST	
PRIMARY ACTORS	Employee
SECONDARY ACTORS	
DESCRIPTION	Employee search information
USER ACTION	IASSIST RESPONSE
1. Employee selects HR bot	2. Chatbot shows an FAQ option for an employee to choose and a text box for a user to search for information
3. Choose FAQ	4. Chatbot sends back information
5. Search for the annual leave policy	6. Return the list of annual leave policy documents
7. Search for payslip	8. Returns a list of payslip information document
9. Search for car and travel allowance policy	10. Returns car and travel allowance policy document
11. Search the organisation's security policy	12. Returns a list of the organisation's security policy document
13. Search for information	14. Chatbot does not contain an answer to the question
ALTERNATE	
Step 13 No answer that matches the question found	<p>The chatbot does not have a definitive answer to the question, the user can check the search results for what has been able to find.</p> <p>Users must also try rephrasing the question using a full sentence, i.e. use more than just a single keyword or two.</p> <p>The question has been saved and will be reviewed so that the chatbot can assist better in the future.</p> <p>Users can also contact HR directly</p>

Table 4.5: Claims iAssist Use Case Narrative

USE CASE NAME: CLAIMS IASSIST	
PRIMARY ACTORS	Employee, Contact Centre Agent
SECONDARY ACTORS	
DESCRIPTION	Employee or Contact Centre Agent search for information
USER ACTION	IASSIST RESPONSE
1. User selects claims bot	2. Chatbot shows an FAQ option for an employee to choose and a text box for a user to search for information
3. Choose FAQ	4. Chatbot sends back information
5. User search for claim policy	6. The bot does not contain a specific answer for this question, it then returns a list of documents that could be related to the search and some options for the related questions the user might want to know
7. User clicks the displayed information about the claim, e.g. when a claim can be rejected	8. The bot returns a list of documents about claim rejection and displays some information on the chat
9. Search the purpose of the claim management framework	10. The bot returns a list of documents about the claim management framework and displays some information on the chat
11. Employee search for who makes business decisions on the claims	12. The bot returns a list of documents about who made a decision on the claim and displays some information on the chat
ALTERNATE	
Step 6 No answer that matches the question found	The chatbot does not have a definitive answer to the question, the user can check the search results for what has been able to find.

Table 4.6: Personal Lines iAssist Use Case Narrative

USE CASE NAME: PERSONAL LINES IASSIST	
PRIMARY ACTORS	Contact Centre Agent
SECONDARY ACTORS	
DESCRIPTION	Contact Centre Agent search for information
USER ACTION	IASSIST RESPONSE
1. User selects personal lines bot	2. Chatbot shows an FAQ option for an employee to choose and a text box for a user to search for information
3. Choose FAQ	4. Chatbot sends back information
5. User search for Underwriting Information	6. The bot does not contain a specific answer for this question, it then returns a list of documents that could be related to the search and some options for the related questions the user might want to know
7. User clicks the displayed information about the underwriting, e.g. underwriting guidelines	8. The bot returns a list of documents about underwriting guidelines and displays some information on the chat
9. Search for policies offered	10. The bot returns a list of documents for the policies that are being offered and displays some information on the chat
11. Agent search for details of allowed risk	12. The bot returns a list of documents about the details of allowed risk and displays some information on the chat
13. Agent searches for items that are allowed to be insured	14. The bot returns a list of documents about the items that are allowed to be insured and displays some information on the chat
ALTERNATE	
Step 6 No answer that matches the question found	The chatbot does not have a definitive answer to the question; the user can check the search results for what has been able to find.

--	--

Table 4.7: Commercial Lines iAssist Use Case Narrative

USE CASE NAME: COMMERCIAL LINES IASSIST	
PRIMARY ACTORS	Contact Centre Agent
SECONDARY ACTORS	
DESCRIPTION	Contact Centre Agent search for information
USER ACTION	IASSIST RESPONSE
1. User selects commercial lines bot	2. Chatbot shows an FAQ option for an employee to choose and a text box for a user to search for information
3. Choose FAQ	4. Chatbot sends back information
5. Agent search for Underwriting Information	6. The bot does not contain a specific answer to this question. It then returns a list of documents that could be related to the search and some options for the related questions the user might want to know
7. User clicks the displayed information about the claim, e.g. underwriting guidelines	8. The bot returns a list of documents about underwriting guidelines and displays some information on the chat
9. Search for policies offered	10. The bot returns a list of documents for the policies that are being offered and displays some information on the chat
11. Agent search for details of allowed risk	12. The bot returns a list of documents about the details of allowed risk and displays some information on the chat
13. Agent searches for items that allowed to be insured	14. The bot returns a list of documents about the items that are allowed to be insured and displays some information on the chat
ALTERNATE	
Step 6 No answer that matches the question found	The chatbot does not have a definitive answer to the question, the user can check the search results for what has been able to find.

--	--

Table 4.8: Policy WhatsApp Use Case Narrative

USE CASE NAME: POLICY WHATSAPP BOT	
PRIMARY ACTORS	Client
SECONDARY ACTORS	
DESCRIPTION	The client starts a chat
USER ACTION	ASSIST RESPONSE
1. Client greet	2. Chatbot shows a welcoming message and the list of services offered by the chatbot (Send my Policy Schedule, Send my Confirmation of Cover, Border Letter, and Home Agent)
3. Choose, and send my confirmation of cover	4. Chatbot prompts the client to enter a preferred method of identification (SA ID number, Namibia ID number, or Passport Number)
5. Client enters the identification	6. Chatbot sends the OTP number to the client's phone number linked to the policy
7. Client enters the OTP number	8. Chatbot sends the document to the client's email that is linked to the policy and responds to the chat
9. Type the question into the text box	10. Chatbot does not have the answer to the question.
11. Client chooses option 4: Home Agent	Chatbot routes the question to the available agent
11. Client enters 0 to exit	12. The chatbot responds with the thank you message
ALTERNATE	
Step 10: Chatbot does not have the answer to the question	Chatbot asks the user to choose the home agent option to route the question to the available agent






4.4 Dataflow Analysis on the Insurance Industry Chatbot

As per the distinction of business process operations, the chatbot is divided into various information streams: refining, transmission, and information storage. In this chapter, the data flow diagram is used to discuss the chatbot business process in detail.

4.4.1 Signs of the Data Flow

In threat modelling analysis, the data flow diagram (DFD) is normally used to imitate the data flow interaction association between chatbot external and internal interactors. The signs of the information flow diagram are shown in Table 4.9.

Table 4.9: Information Flow Signs

Signs	Sign Names	Description
	User interaction	User input to the chatbot
	Process	Information manipulation
	Information Storage	Permanent and temporal information storage
	Information Flow	Shows information flow from information stores, processes, or interactors
	Boundaries	The system, physical, address space, or trust boundary.

4.5 Analysis of the Insurance Chatbot

4.5.1 iAssist Chatbot

iAssist chatbot is used to assist employees with access to frequently asked questions, as well as in FAQ through cognitive or intelligent search. It is beyond just the chat; it is also a search bot and document web document viewer. It is exposed to the internal internet environment which is an intranet.

The user interacts with the chatbot mainly through:

- User login (1.0): for a user to access the chatbot operations needs to log in to the system and the user access the information based on the user's role.

- Claim (2.0): anything that has to do with claim information or documents, such as adding basic access and waiver, rules regarding the claim, etc.
- Personal Lines (3.0): This can be initiated by agent users, personal line functionality provide the corresponding underwriting information, such as underwriting guidelines for personal lines.
- Commercial Lines (4.0): This can be initiated by agent users, personal line functionality provide the corresponding underwriting information, such as underwriting guidelines for commercial lines.
- Human Resource(5.0): This can be accessed by all employees across the organisation, it provides HR-related information and documents.

The first level data flow diagram decomposition of these business process operations is shown in Figure 4.3

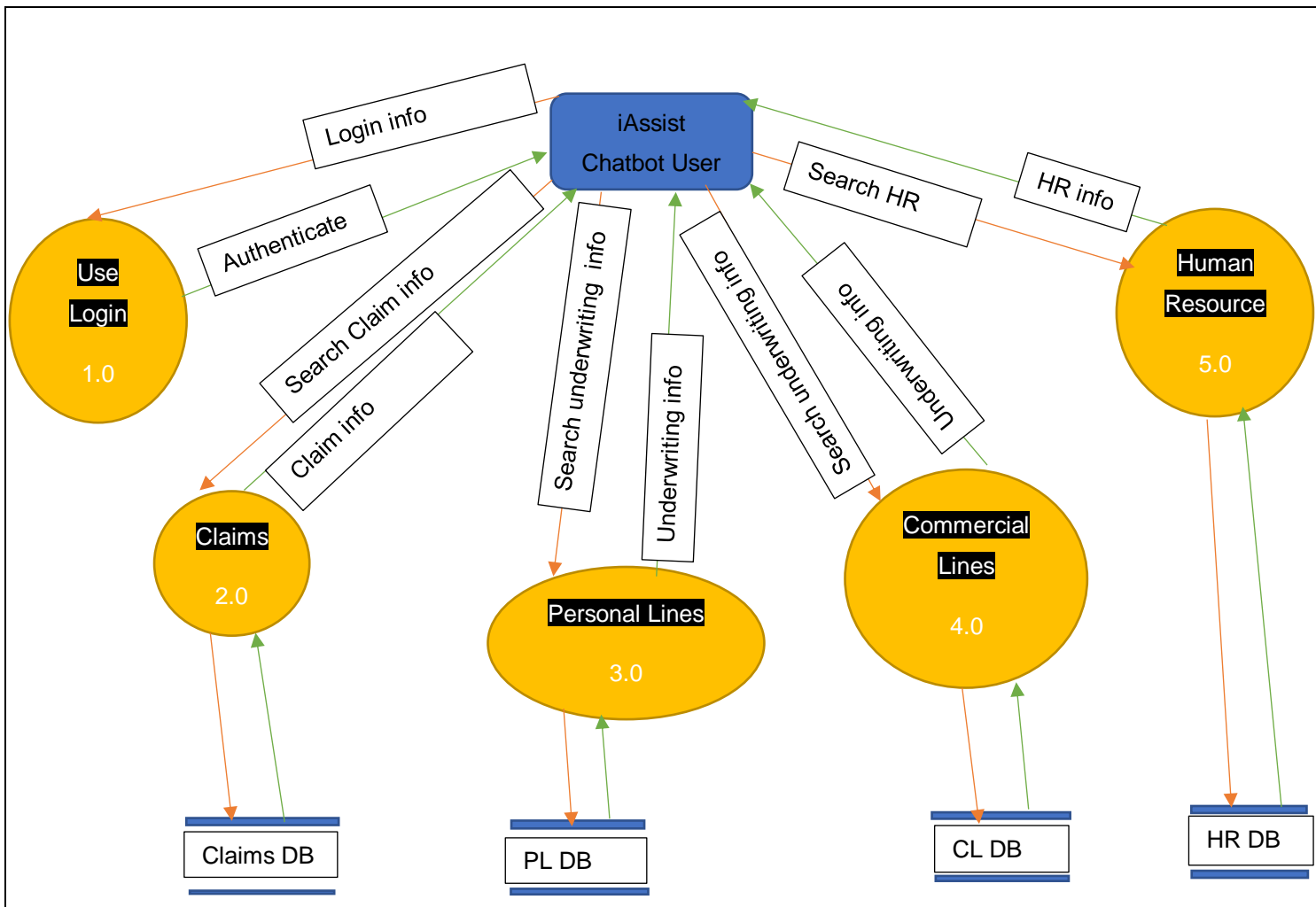


Figure 4.3: Critical Business Data Flow of iAssist Chatbot

Figures 4.4 to Figure 4.8 give the second level of data flow diagram decomposition of the five different business operations (User Login, Claims, Personal Lines, Commercial Lines, Human Resource) of the iAssist chatbot.

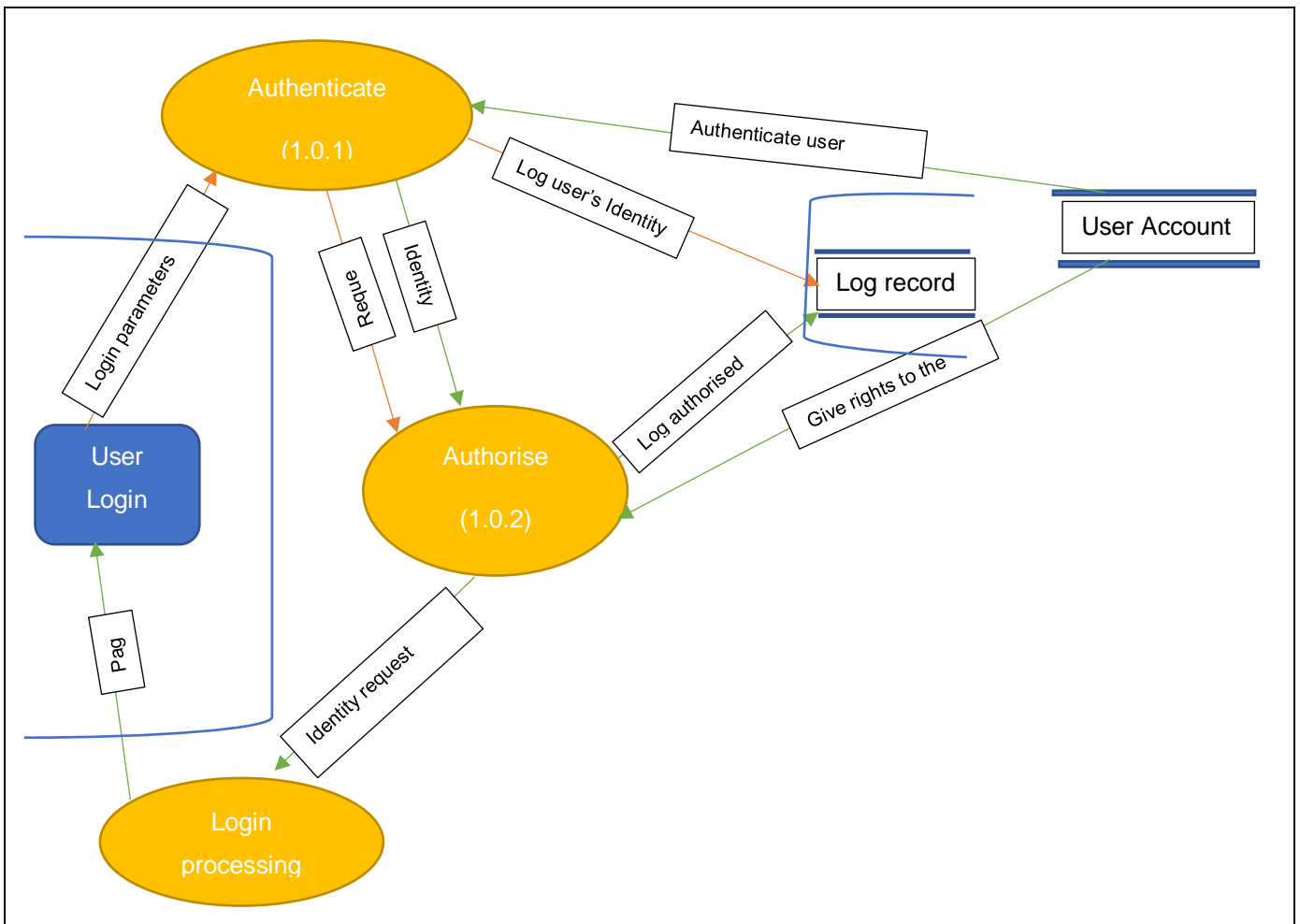


Figure 4.4: Login Process

Figure 4.4 depicts the login process. Before the user is given access to a chatbot, the user needs to log in first with a username and password for authentication. Authentication is done through the user account database. Once the user is authorised then the user accesses the chatbot based on the user's role. All the interactions are stored on the log file for auditing purposes.

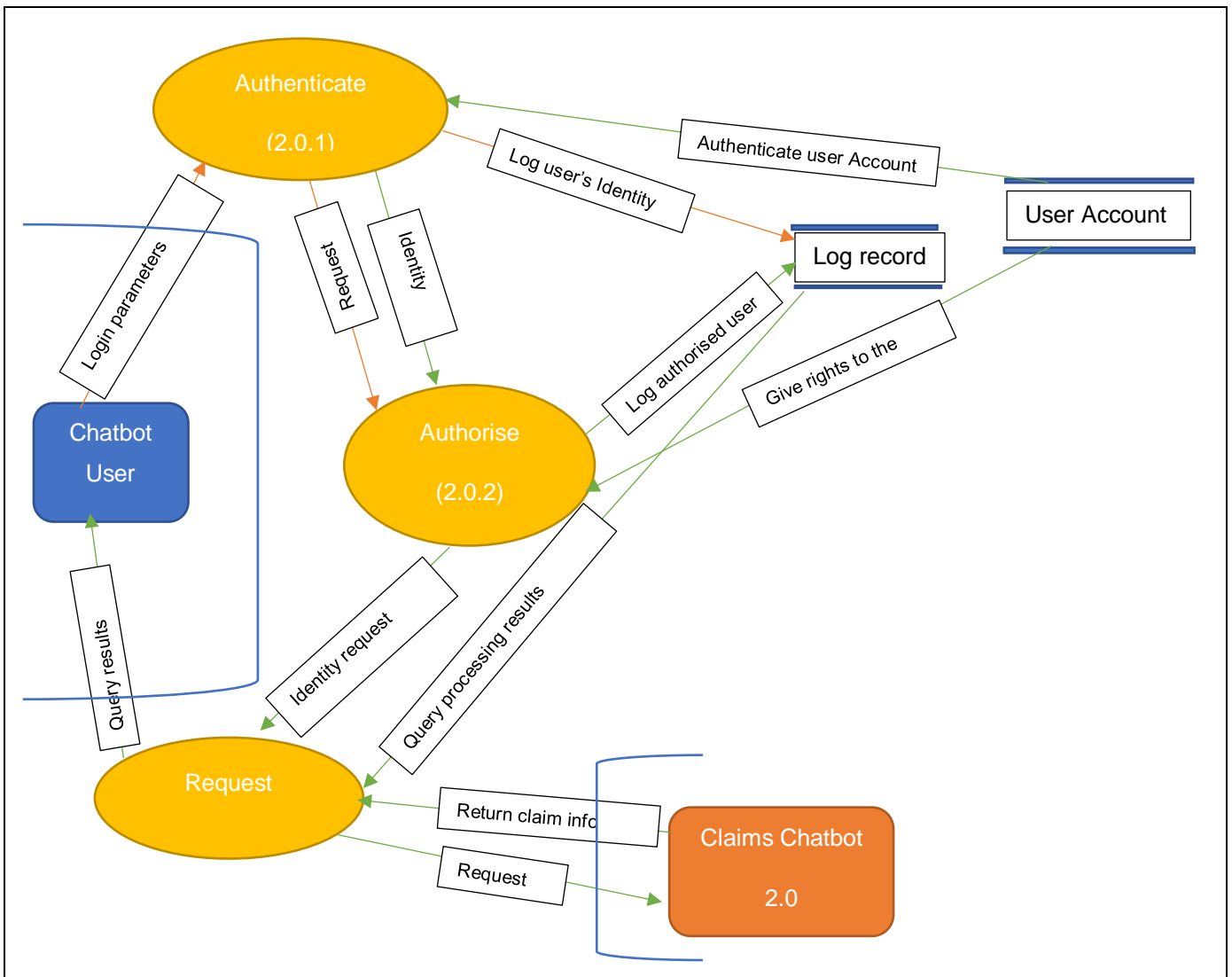


Figure 4.5: Claim Request

Figure 4.5 depicts when the user has already been given rights to access the Claims chatbot. Then the user requests information and asks FAQ (frequently asked questions) related to the claim. All the interactions including query processing results are stored in the log file for auditing purposes.

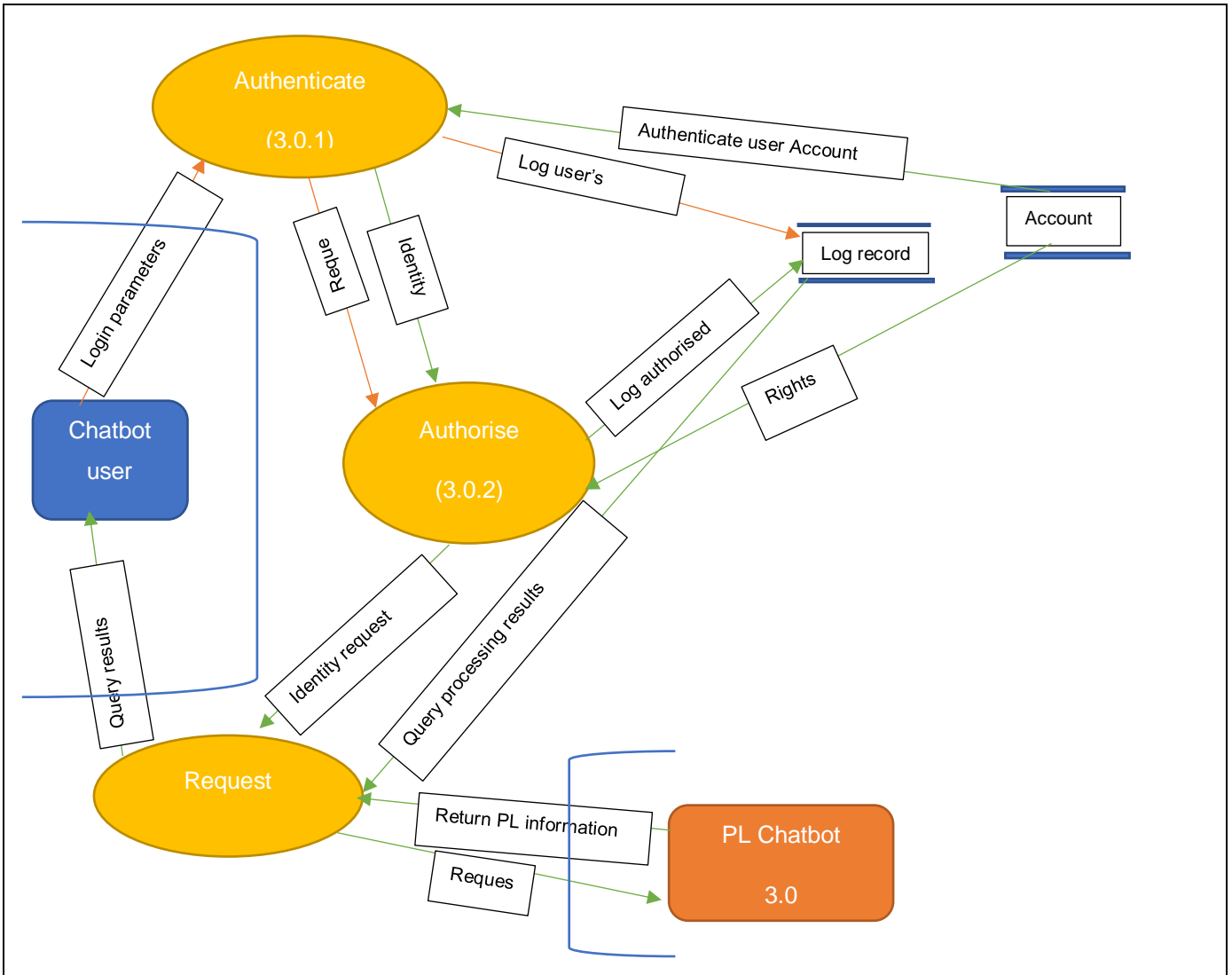


Figure 4.6: Personal Lines Request

Figure 4.6 shows when the user has been given rights to access the Personal Lines chatbot. The user request information and ask FAQ related to the Personal Lines queries. All the interactions with the chatbot including query processing results are stored in the log file for auditing purposes.

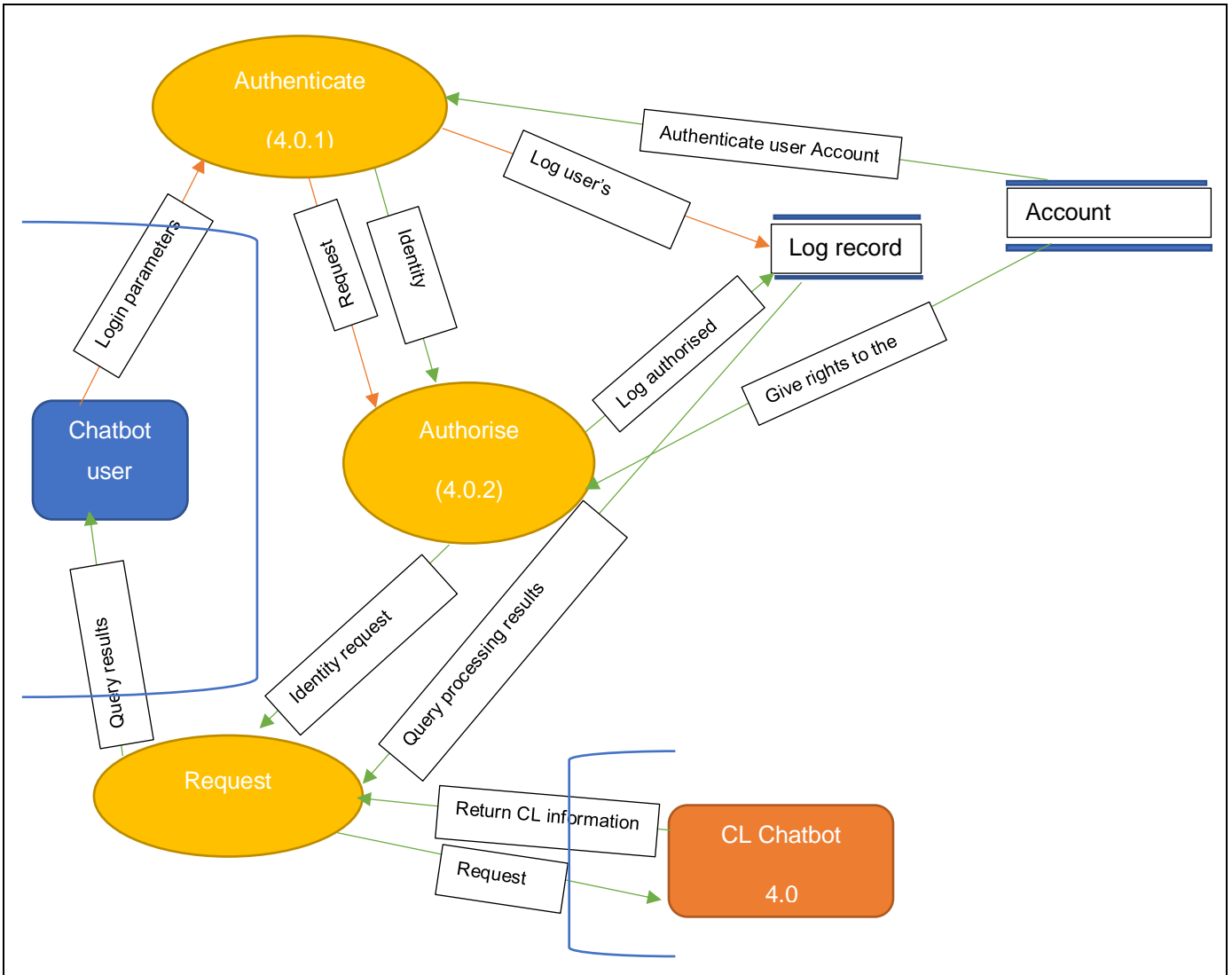


Figure 4.7: Commercial Lines Request

Figure 4.7 shows when the user has been given rights to access the Commercial Lines chatbot. The user request information and ask FAQ related to the Commercial Lines queries. All the interactions with the chatbot including query processing results are stored in the log file for auditing purposes.

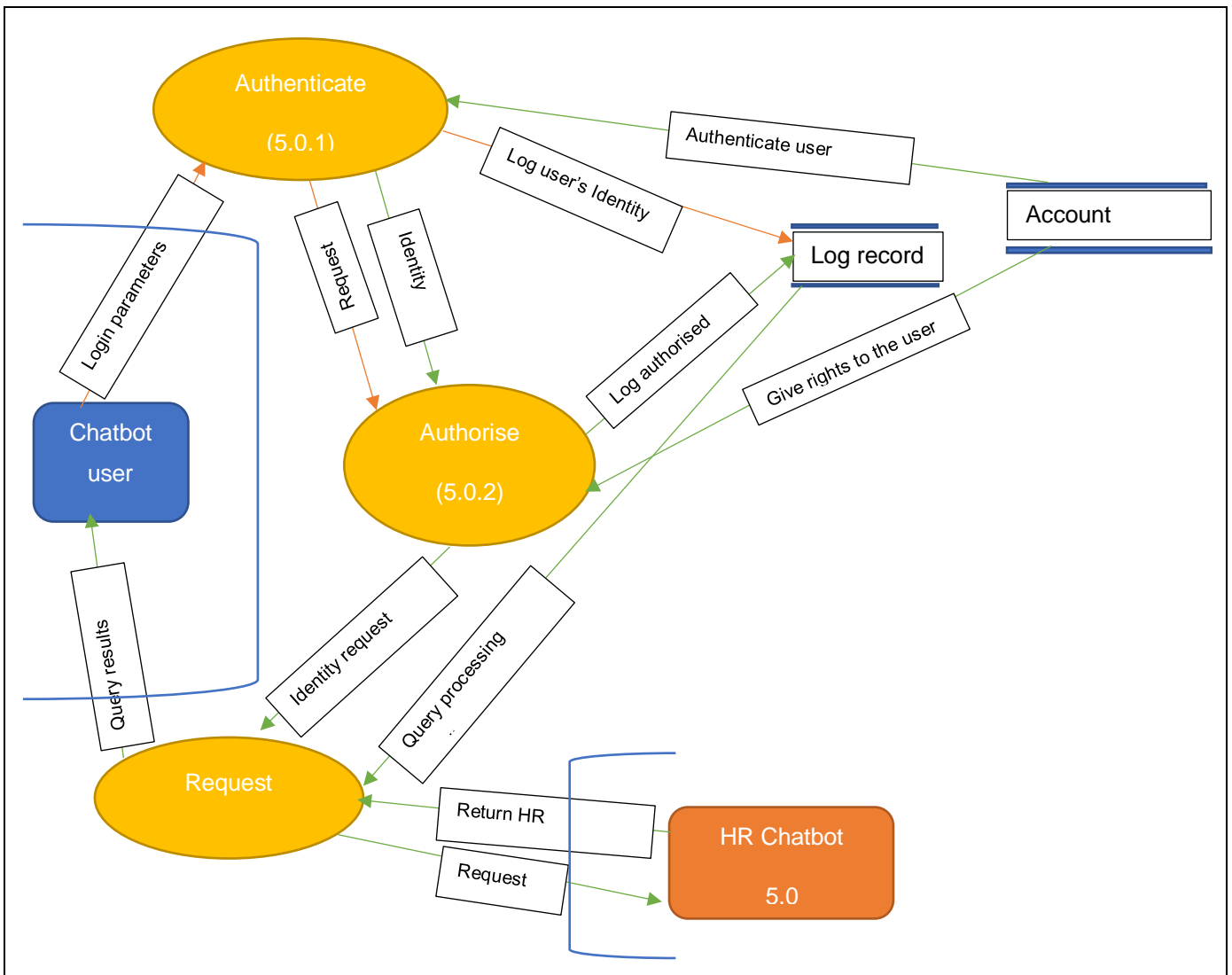


Figure 4.8: Human Resource Request

Figure 4.8 shows when the user has been given rights to access the Human Resource chatbot. The user request information and ask FAQ related to the Human Resource queries. All the interactions with the chatbot including query processing results are stored in the log file for auditing purposes.

4.5.2 WhatsApp Chatbot

WhatsApp Insurance chatbot manages a relationship with clients effectively. It cut out the middleman manual intervention of customers having to speak directly with consultants by providing a self-service. It provides easy and fast access to information, consistency of response, is easy to use, and provides 24/7 access. The user interactors with the chatbot mainly include User verification (1.0): for a user to access the chatbot operations, they need to provide an ID or passport number, the OTP is sent to the user's registered mobile number for verification and all the requests are sent to the client's email.

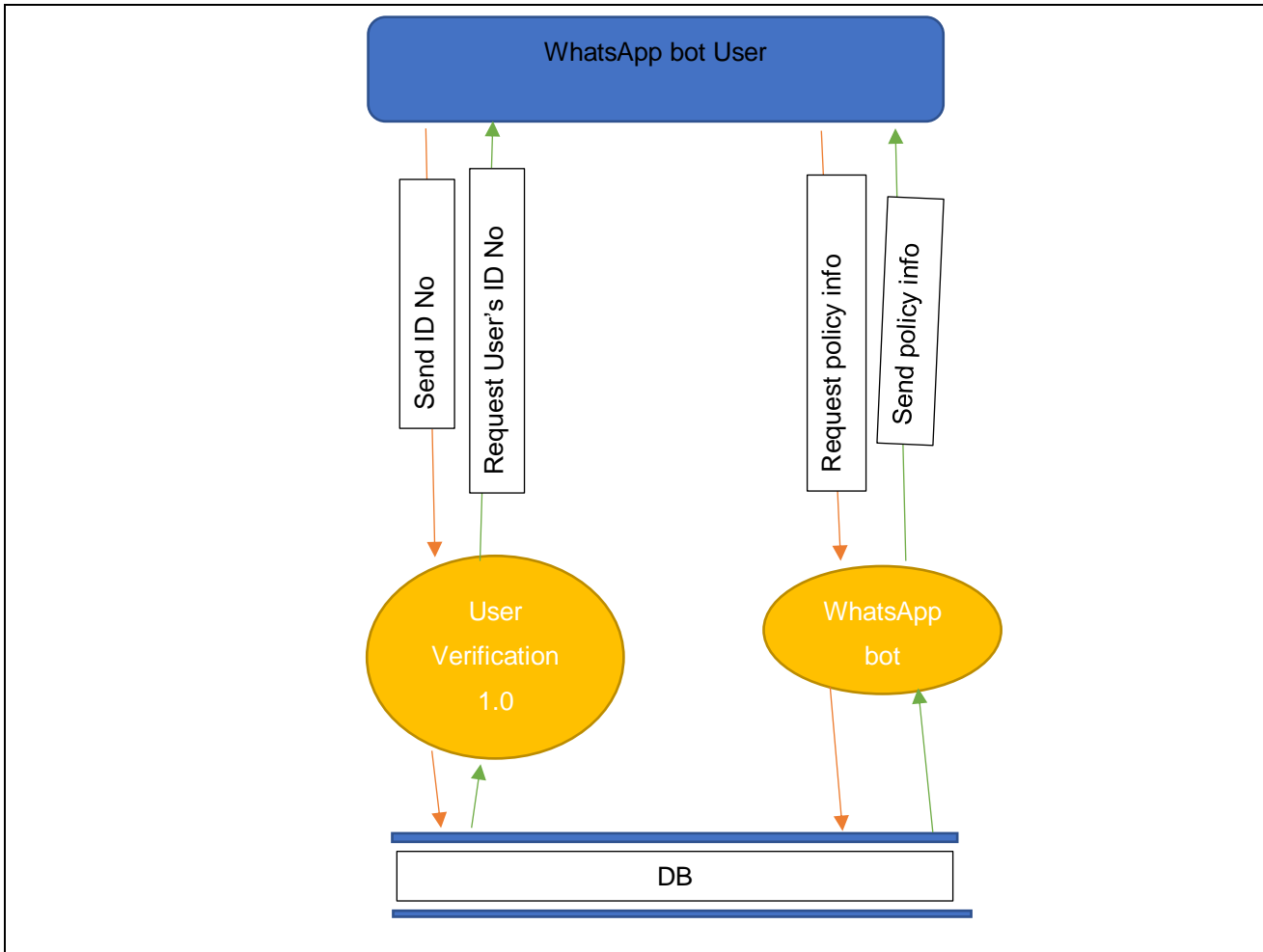


Figure 4.9: WhatsApp Business Data Flow

Figure 4.9 gives the second level of the WhatsApp data flow diagram decomposition of the above business operations, respectively.

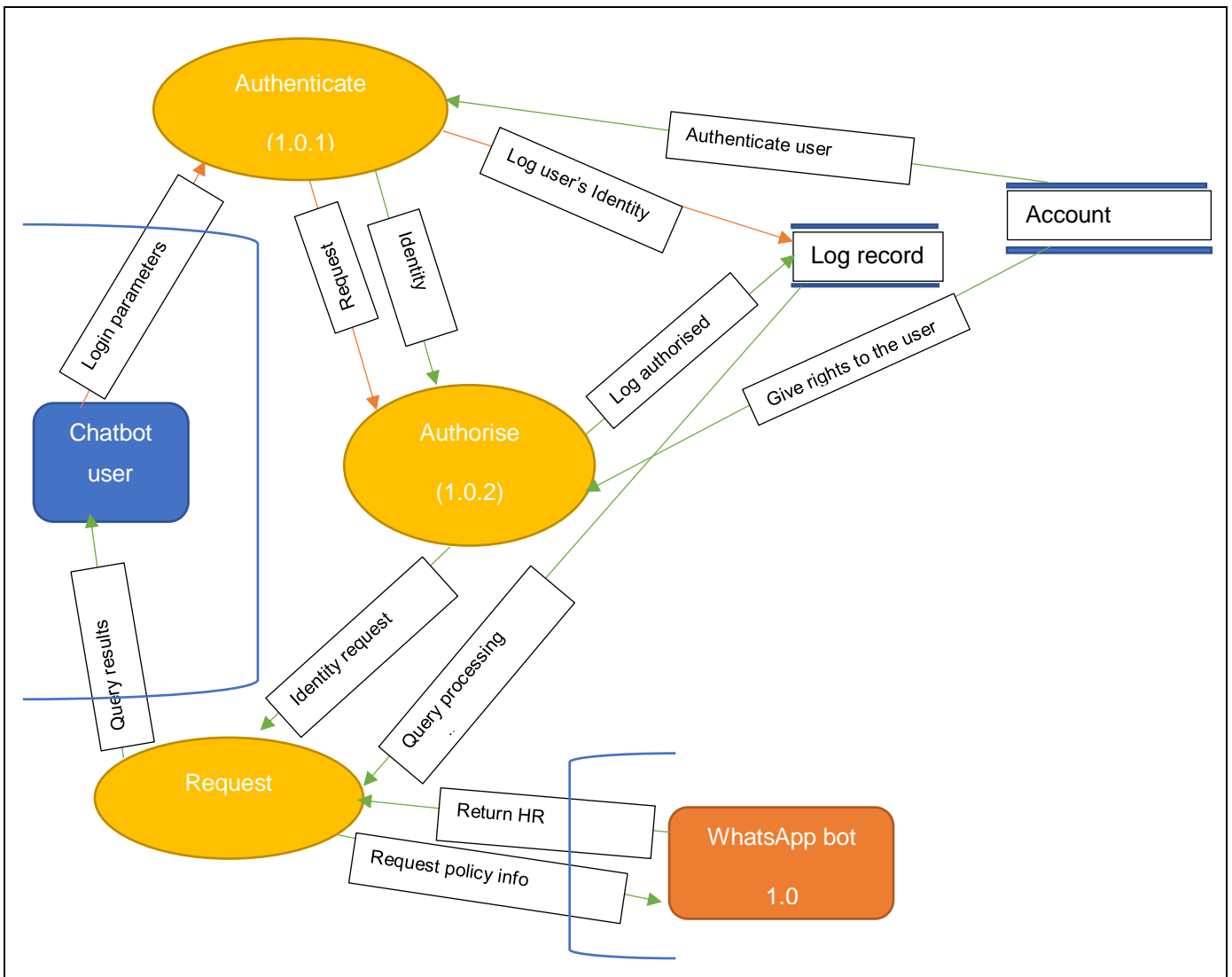


Figure 4.10: WhatsApp Bot User Interaction

Figure 4.10 shows when the user has been given rights to access the WhatsApp chatbot. The user request information and ask FAQ related to the policy queries. All the interactions with the chatbot including query processing results are stored in the log file for auditing purposes.

4.6 Chapter Summary

This chapter gives a description of the case study and how the data was collected from the case study and analysed. The profile of the study's participants is explained. Findings from the analysis of participants' responses are presented. Based on objectives one and two of the study, which are the insurance chatbot use cases, a list of observations was identified from the collected data.

CHAPTER FIVE

THREAT MODEL DEVELOPMENT

This section presents the threat models developed for data security in insurance chatbots based on STRIDE modelling.

5.1 Modelling Security based on STRIDE

The STRIDE model derives from the following six threat groups: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege. The component of STRIDE is explained further as follows:

- i. Spoofing- This threat happens when the attacker illegally accesses and uses another user's credentials. Impersonating something or someone else.
- ii. Tampering- This threat happens when attackers aim to maliciously change/modify data.
- iii. Repudiation- This threat happens when an attacker aims to perform an illegal operation in a system
- iv. Information Disclosure- hackers steal confidential information
- v. Denial of Service- This threat happens when an attacker aims to deny access to valid users.
- vi. Elevation of Privilege- This threat happens when an attacker aims to gain privileged access.

Considering the six types of threats covered by STRIDE, for every element in the Chatbot Data Flow Diagram, STRIDE modelling was used to identify the security threats pertaining to the chatbot.

The results of the STRIDE modelling are shown in Tables 5.1 – 5.6. The threat model diagrams examine whether each chatbot and its related asset data is vulnerable to any of the security threats in STRIDE. For each aspect of STRIDE, the scenario description of the threat in terms of security goals, security threats, and security vulnerability is presented in a table, together with the threat model (diagram), and a textual description of the threat model.

Tables 5.1 – 5.6 show the security goals, security threats, and security vulnerabilities for each chatbot in the organisation. The symbol (√) means that the identified security goals, security threats, and security vulnerability pertain to a specific chatbot, while the symbol (X) means the identified security goals, security threats, and security vulnerability do not pertain to a specific chatbot.

The threat model diagrams show the attack possibilities each chatbot might encounter and the mitigation or defence for the attack. The diagrams were developed by using an Attack Defence tool called ADTool (Arias, Petrucci, Masko, Penczek & Sidoruk, 2022). The threat model is used to show the actions of an attacker trying to compromise the system and the possible counteractions of a defender trying to protect the system. It also helps with the qualitative security analysis, using attack–defence trees. The ADTool represents the attacks with red circles and the defence with green boxes.

5.2 Notations Used for Representation of the Insurance Chatbots

For easier representation, the different insurance chatbots were represented with symbols as follows:

- i. Claims bot (AI Bot) – A1
- ii. Personal Lines bot (AI Bot) – A2
- iii. Commercial Lines bot (AI Bot) – A3
- iv. Human Resource bot (AI Bot)- A4
- v. WhatsApp bot (Declarative / Human Bot) – A5

5.3 Spoofing

5.3.1 Scenario Description of Spoofing

Table 5.1 shows the security goal, the source of security threats for spoofing, and how insurance chatbots are vulnerable to spoofing. Chatbots A1 to A5 are not only vulnerable to social engineering attacks and chatbot A5 is only not vulnerable to errors in business logic and business process.

Table 5.1: Spoofing

	A1	A2	A3	A4	A5
Security Goal					
To provide for the confidentiality, integrity, and availability of the organisation's information assets by ensuring that users are authenticated and authorised to access these information assets	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors	√	√	√	√	√
Security Vulnerabilities					
Errors in business logic and business processes	√	√	√	√	X
Social engineering attacks	X	X	X	X	√
Interception of communication	√	√	√	√	√
Unauthorised access	√	√	√	√	√
Forms: <ul style="list-style-type: none"> • Information leakage • Insecure authentication • Brute force attacks 					

<ul style="list-style-type: none"> Exploiting unencrypted / poorly encrypted 					
Attack through bad design or mis-configuration Forms: <ul style="list-style-type: none"> Vulnerabilities in authentication Poor implementation of encryption protocols 	√	√	√	√	√
phishing attacks Forms: <ul style="list-style-type: none"> use phishing tools send spam emails Randomly targeting many individuals or organizations 	√	√	√	√	√
Insider threats	√	√	√	√	√
Manipulation of employees by attackers Forms: <ul style="list-style-type: none"> disclosing private or confidential information giving unauthorised physical access social engineering techniques Influence an employee or legitimate user Bribe employees Blackmail employees 	√	√	√	√	√
Possible Security Threats					
An attacker can exploit insecure default configuration	√	√	√	√	√
An attacker can access the system from an unauthorised network	√	√	√	√	√
An attacker can execute administrator functions	√	√	√	√	√
An attacker can gain unauthorised access	√	√	√	√	√
Insecure infrastructure	√	√	√	√	√

5.3.2 Threat Model for Spoofing

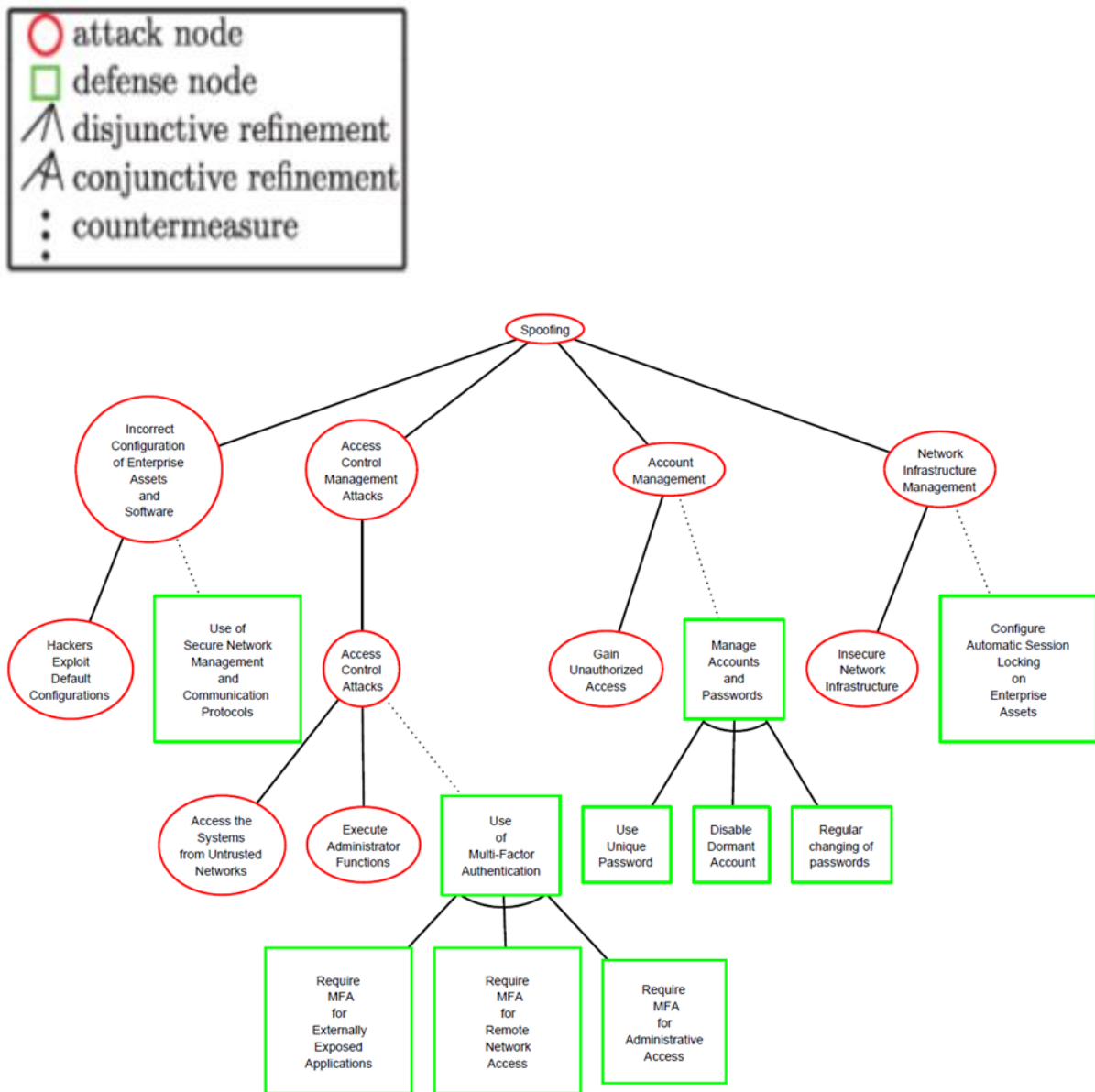


Figure 5.1: Spoofing Attack and Defence

5.3.3 Description of the Threat Model for Spoofing

According to Figure 5.1, spoofing attacks can be through the incorrect configuration of enterprise assets, account management, access control management, and network infrastructure.

- i. The mitigation for incorrect configuration of enterprise assets and its child (exploit default configuration) is the use of secure network management and communication protocol.
- ii. The mitigation for access control and its child: access control attacks which also have two children (access the systems from untrusted networks and execute administrator functions)

is the use of multi-factor authentication which could be MFA for externally exposed applications, MFA for remote network access and MFA for administrative access.

- iii. The mitigation for account management and its child (gain unauthorised access) is to manage account passwords which could be through the use of a unique password, disabling the dormant account and regular dormant account.
- iv. The mitigation for network infrastructure management and its child (Insecure network infrastructure) is to configure automatic session locking on enterprise assets.

5.4 Tampering with Data

5.4.1 Scenario Description of Tampering with Data

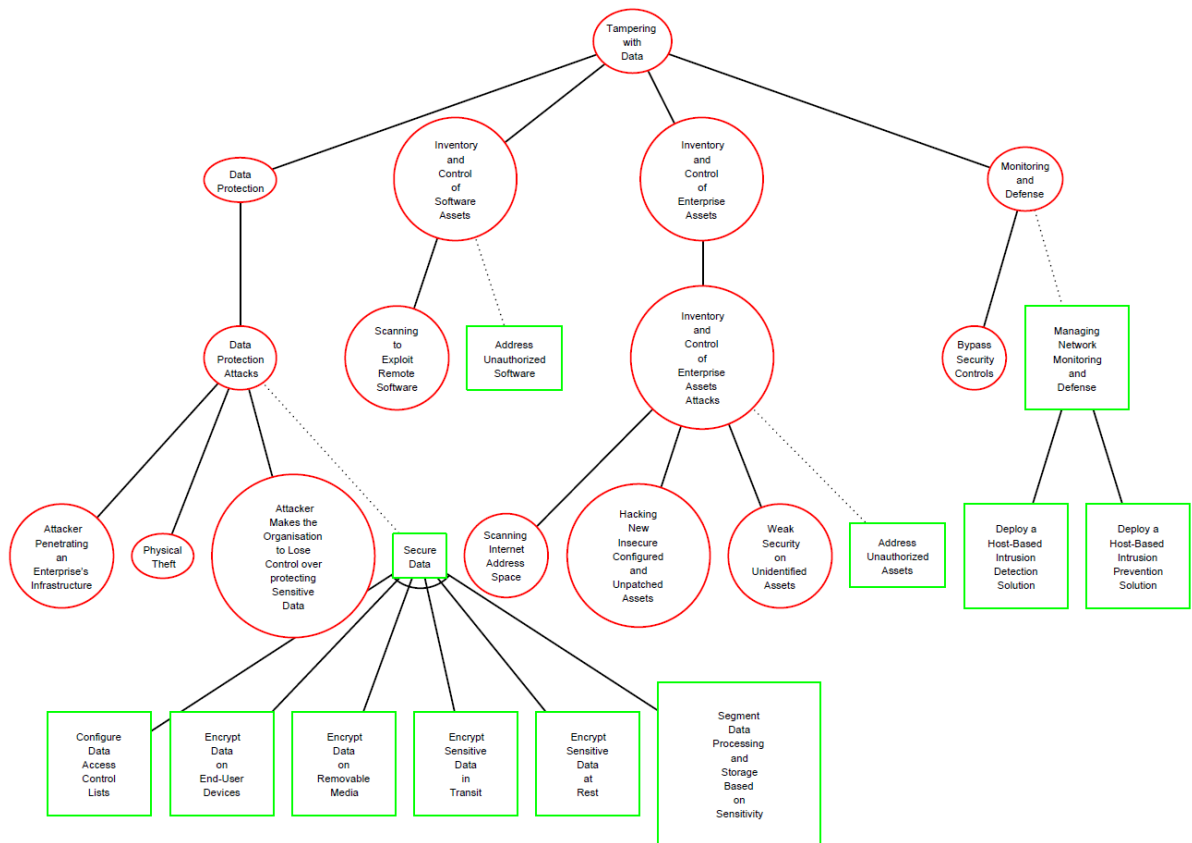
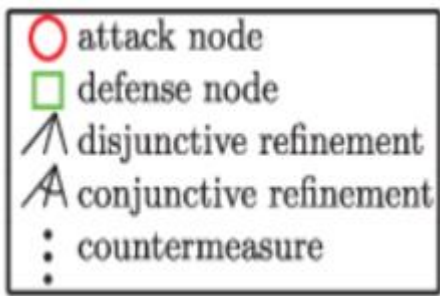
Table 5.2 shows the security goal, the source of security threats for tampering with the data component, and how insurance chatbots are vulnerable to this component.

Table 5.2: Tampering with Data

	A1	A2	A3	A4	A5
Security Goal					
To provide for the integrity and confidentiality of the organisation's information assets by ensuring that information assets are accurately and authentically represented	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors	√	√	√	√	√
Security Vulnerabilities					
Gain access to information assets in transit (network sniffing) Forms: <ul style="list-style-type: none"> • Man-in-the-middle attacks • Domain Name System Hijacking • Communication not encrypted or weakly encrypted • Real-time traffic modification 	√	√	√	√	√
Information compromise forms: <ul style="list-style-type: none"> • Phishing 	√	√	√	√	√

<ul style="list-style-type: none"> Pivot off a foothold in a third party's information applications 					
Possible Security Threats					
Attacker penetrating an enterprise's infrastructure	√	√	√	√	√
An attacker can perform a physical attack in which a physical theft occurs (e.g. stealing portable devices, etc.)	X	X	X	X	X
Physical theft	√	√	√	√	√
An attacker can make organisations lose control over protecting sensitive data	√	√	√	√	√
Hackers scan for vulnerabilities to exploit the remote software	√	√	√	√	√
Scanning internet address space	√	√	√	√	√
Hacking new insecure configurations and unpatched assets	√	√	√	√	√
Weak security on unidentified assets	√	√	√	√	√
An attacker can bypass security controls	√	√	√	√	√
Lack of adequate logging and regular log review in applications	√	√	√	√	√
Poor authentication and incorrect testing	√	√	√	√	√
An attacker can exploit insecure default configuration	√	√	√	√	√
Insecure network infrastructure	√	√	√	√	√
An attacker can make changes to configurations	√	√	√	√	√

5.4.2 Threat Model for Tampering with Data



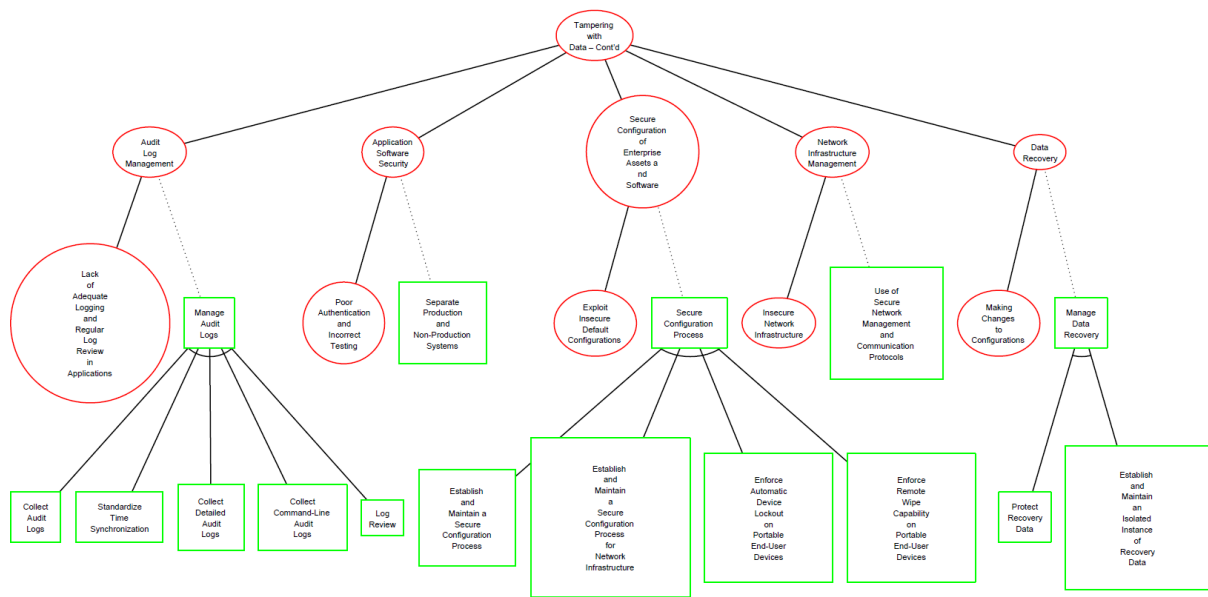


Figure 5.2 Tampering with Data Attack and Defence

5.4.3 Description of the Threat Model for Tampering with Data

According to Table 5.2, tampering with data attacks can be through data protection, inventory and control of software assets, inventory and control of enterprise assets, audit log file management, monitoring and defence, application software security, secure configuration of enterprise assets and software, network infrastructure management, and data recovery.

- i. The mitigation for data protection and its child: data protection attacks which also have two children (attacker penetrating an enterprise's infrastructure, physical theft, the attacker makes the organisation lose control over protecting sensitive data) is the use of secure data which could be configuring data access control lists, encrypt data on end-user devices, encrypt data on removable media, encrypt sensitive data in transit, encrypt sensitive data at rest, segment data processing and storage based on sensitivity.
- ii. The mitigation for inventory and control of software assets and their child (scanning for vulnerabilities to exploit remote software) is to address unauthorised software.
- iii. The mitigation for inventory and control of enterprise assets and its child: inventory and control of enterprise assets attacks which also have two children (scanning internet address space, hacking new insecure configured and unpatched assets, weak security on unidentified assets) are used to address unauthorised assets.
- iv. The mitigation for monitoring and defence and its child (bypass security controls) is through managing network monitoring and defence, which includes deploying a host-based intrusion detection solution and deploying a host-based intrusion prevention solution.
- v. The mitigation for audit log management and its child (lack of adequate logging and regular

- vi. log review in applications) is through managing audit logs, which include collecting auditing logs, standardising time synchronisation, collecting detailed audit logs, command-line audit logs, and log review.
- vii. The mitigation application software security and its child (poor authentication and incorrect testing) to separate production and non-production systems.
- viii. The mitigation for secure configuration on enterprise assets and software and its child (exploit insecure default configurations) is through a secure configuration process which includes establishing and maintaining a secure configuration process, establishing and maintaining a secure configuration process for network infrastructure, automatic device lockout on portable end-user devices, and enforcing remote wipe capability on portable end-user devices.
- ix. The mitigation network infrastructure management and its child (insecure network infrastructure) is to use secure network management and communication protocols.
- x. The mitigation of data recovery and its child (making changes to configurations) is through managing data recovery, which includes protecting recovery data and establishing and maintaining an isolated instance of recovery data.

5.5 Repudiation

5.5.1 Scenario Description of Repudiation

Table 5.3 shows the security goal, the source of security threats for the repudiation component, and how insurance chatbots are vulnerable to this component.

Table 5.3: Repudiation

	A1	A2	A3	A4	A5
Security Goal					
To provide for the integrity of the organisation's information assets by ensuring that information assets are accurately and authentically represented and that there is a robust audit trail proving who performed each action being audited as it relates to information assets.	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors	√	√	√	√	√
Security Vulnerabilities					
Impersonating real users	√	√	√	√	√

Forms: <ul style="list-style-type: none"> • Accessing Personal and Private Information • Performing Transactions 					
Possible Security Threats					
An attacker can exploit insecure default configuration	√	√	√	√	√
Attacker penetrating an enterprise's (infrastructure)	√	√	√	√	√
An attacker can perform a physical attack in which a physical theft occurs (e.g. stealing portable devices, etc.)	X	X	X	X	X
An attacker can make organisations lose control over protecting sensitive data	√	√	√	√	√
Lack of adequate logging and regular log review in applications	√	√	√	√	√
Insecure network infrastructure	√	√	√	√	√

5.5.2 Threat Model for Repudiation

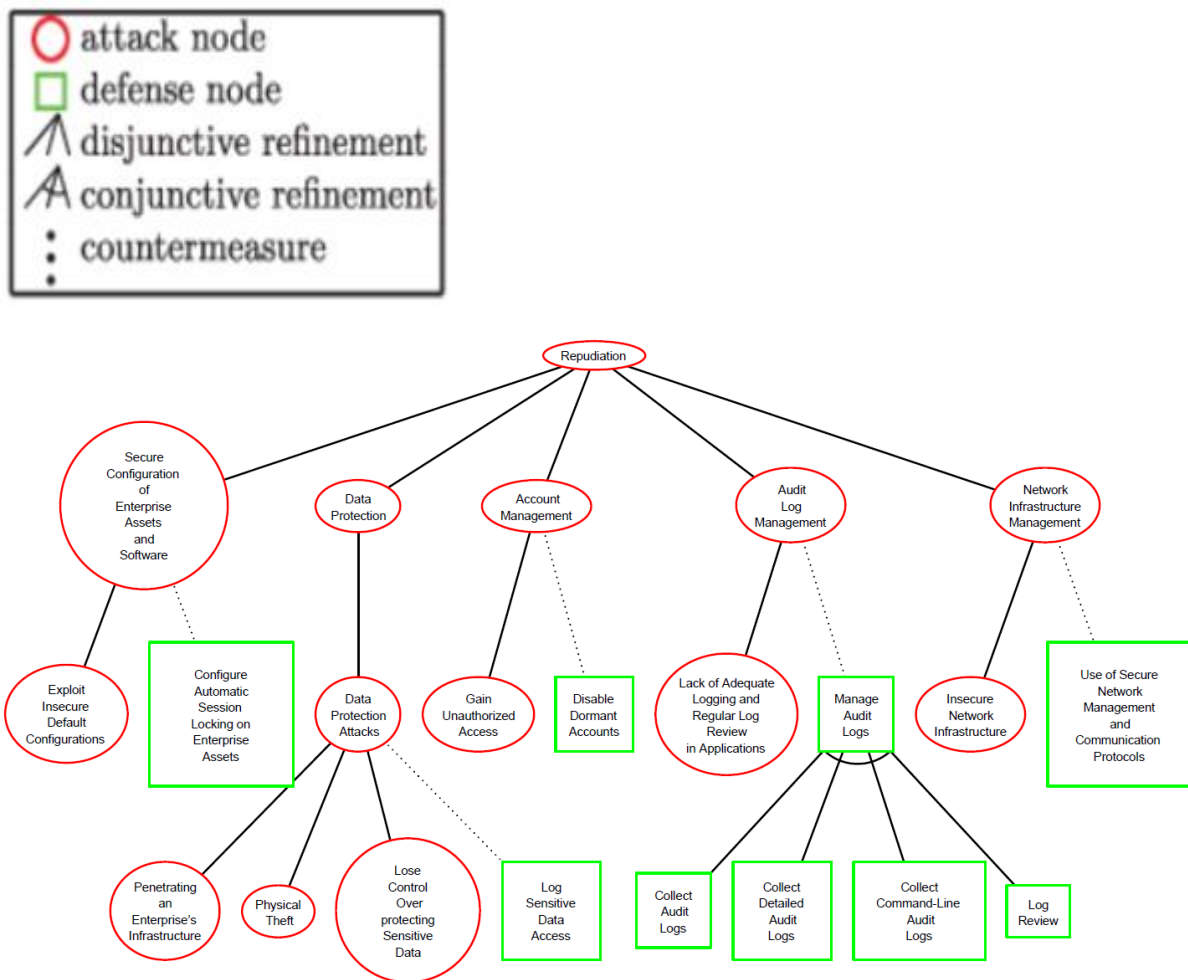


Figure 5.3: Repudiation Attack and Defence

5.5.3 Description of the Threat Model for Repudiation

According to Figure 5.3, repudiation attacks can be through the secure configuration of enterprise assets and software, data protection, account management, audit log file management, and network infrastructure management.

- i. The mitigation for secure configuration on enterprise assets and software and its child (exploit insecure default configurations) is to configure automatic session locking on enterprise assets.
- ii. The mitigation for data protection and its child: data protection attacks which have also two children (attacker penetrating an enterprise's infrastructure, physical theft, the attacker making the organisation lose control over protecting sensitive data) is to log sensitive data access.
- iii. The mitigation for account management and its child (gain unauthorized access) is to disable dormant accounts.

- iv. The mitigation for audit log management and its child (lack of adequate logging and regular log review in applications) is through managing audit logs which includes collecting audit logs, collecting detailed audit logs, collecting command-line audit logs, and log review.
- v. The mitigation for network infrastructure management and its child (Insecure network infrastructure) is to use secure network management and communication protocols.

5.6 Information Disclosure

5.6.1 Scenario Description of Information Disclosure

Table 5.4 shows the security goal, the source of security threats for the information disclosure component, and how insurance chatbots are vulnerable to this component.

Table 5.4: Information Disclosure

	A1	A2	A3	A4	A5
Security Goal					
To provide for the confidentiality of the organisation's information assets by ensuring that information assets are only accessible to authenticated and authorised individuals.	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors	√	√	√	√	√
Security Vulnerabilities					
The insecure disposal of assets. Forms: <ul style="list-style-type: none"> • Data deletion • Insecure disposal of data assets related to the environment 	√	√	√	√	√
Possible Security Threats					
An attacker can exploit insecure default configuration	√	√	√	√	√
Attacker penetrating an enterprise's infostructure (infostructure)	√	√	√	√	√

An attacker can perform a physical attack in which a physical theft occurs (e.g. stealing portable devices, etc.)	√	√	√	√	√
An attacker can make organizations lose control over protecting sensitive data	√	√	√	√	√
Hackers scan for vulnerabilities to exploit the remote software	√	√	√	√	√
Ransomware attacker from the third party	√	√	√	√	√

5.6.2 Threat Model for Information Disclosure

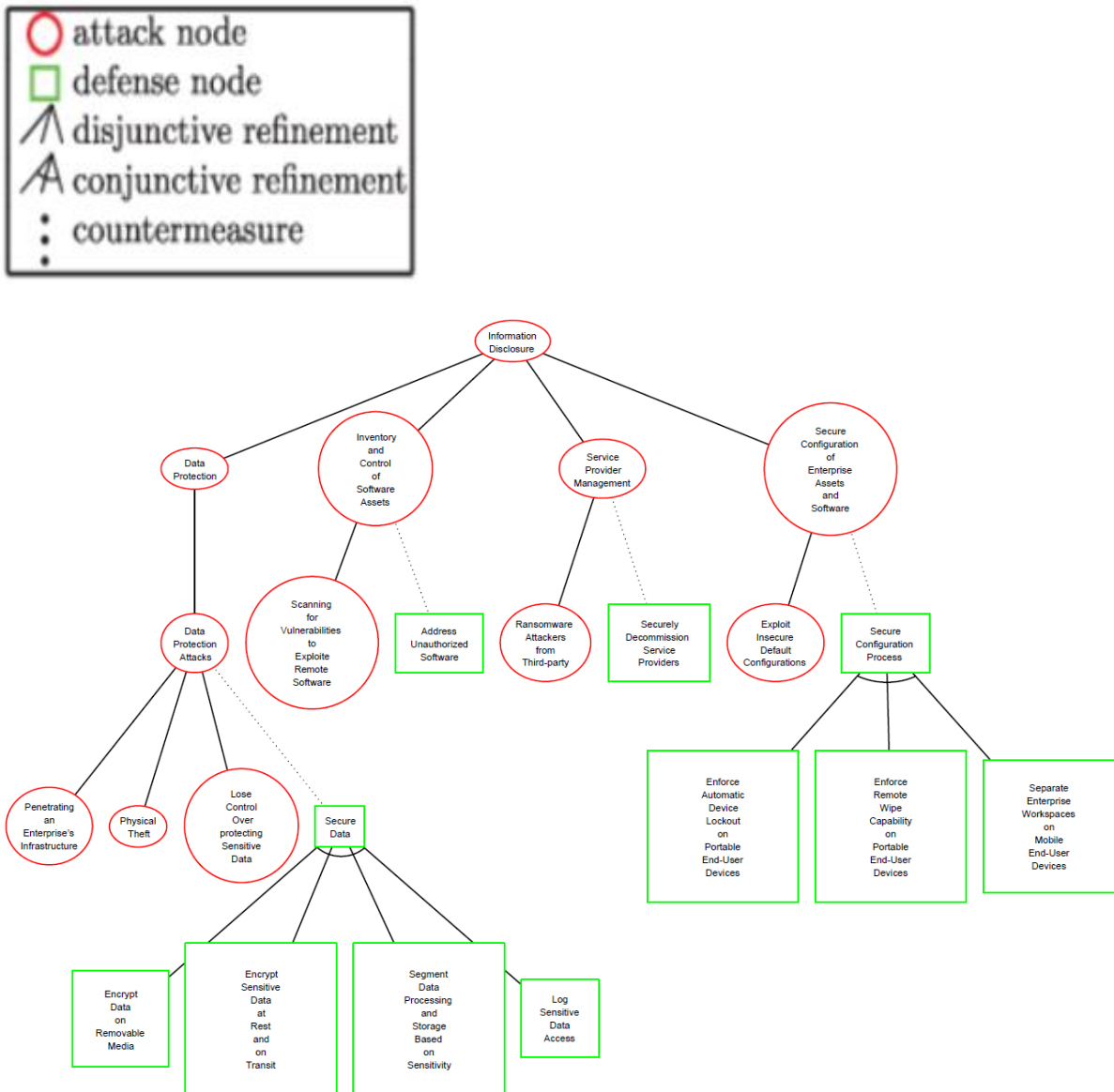


Figure 5.4 Information Disclosure Attack and Defence

5.6.3 Description of the Threat Model for Information Disclosure

According to Figure 5.4, Information Disclosure attacks can be through data protection, inventory and control of software assets, service provider management, and secure configuration of enterprise assets and software.

- i. The mitigation for data protection and its child: data protection attacks which also have two children (attacker penetrating an enterprise’s infrastructure, physical theft, attacker making the organisation lose control over protecting sensitive data); the use of secure data which could be to configure data access control lists, encrypt data on end-User devices, encrypt data on removable media, encrypt sensitive data in transit, encrypt sensitive data at rest, segment data processing and storage based on sensitivity.
- ii. The mitigation for inventory and control of software assets and its child (scanning for vulnerabilities to exploit remote software) is to address unauthorised software.
- iii. The mitigation for service provider management and its child (ransomware attackers from third-party) is securely decommissioning service providers.
- iv. The mitigation for secure configuration on enterprise assets and software and its child (exploit insecure default configurations) is through a secure configuration process, which includes establishing and maintaining a secure configuration process, establishing and maintaining a secure configuration process for network infrastructure, automatic device lockout on portable end-user devices, and enforcing remote wipe capability on portable end-user devices.

5.7 Denial of Service

5.7.1 Scenario Description of Denial of Service

Table 5.5 shows the security goal, the source of security threats for the denial of service component, and how insurance chatbots are vulnerable to this component.

Table 5.5: Denial of Service

	A1	A2	A3	A4	A5
Security Goal					
To ensure that organisational service level agreements and availability objectives are met. This includes ensuring that business-critical processes are available to service client and business partner transactions.	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-	√	√	√	√	√

states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors					
Security Vulnerabilities					
Impair the availability or performance of the organisation's applications Forms: <ul style="list-style-type: none"> • Use of publicly available tools (a single source, single host, and limited network bandwidth) • Utilisation of custom tools • Amount of network bandwidth • Interfering with wireless communications 	√	√	√	√	√
Introduce malware into the application Forms: <ul style="list-style-type: none"> • Creation of custom-written malware • Individual malware attacks (through their mobile devices, influencing them to use infected websites: infected portable storage devices) • Use of rootkits or anti-forensics methods 	√	√	√	√	√
Badly designed network architecture Forms: <ul style="list-style-type: none"> • Badly designed network architecture • Insecure or vulnerable Internet connections • Poor filtering on Internet or internal network connections • Absence of segregation of critical application functions 	√	√	√	√	√
Physical damage or tampering with the company's data application	√	√	√	√	√
Stealing of physical infrastructure					
Possible Security Threats					
Lack of adequate logging and regular log review in applications	√	√	√	√	√
An attacker can make changes to configurations	√	√	√	√	√
Insecure network infrastructure	√	√	√	√	√

5.7.2 Threat Model for Denial of Service

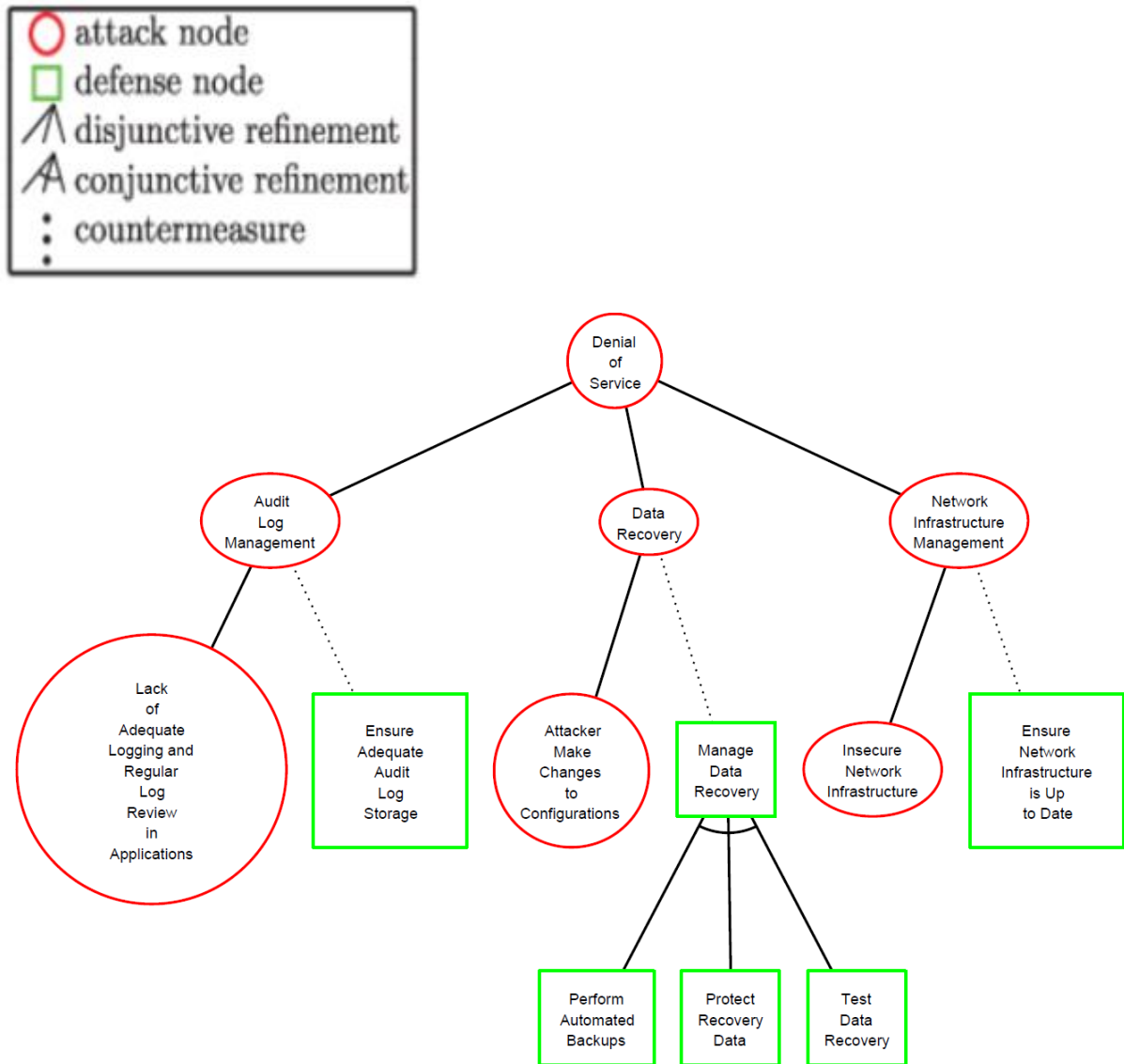


Figure 5.5: Denial of Service Attack and Defence

5.7.3 Description of the Threat Model for Denial of Service

According to Figure 5.5, denial of Service attacks can be through audit log management, data recovery, and network infrastructure management.

- i. The mitigation for audit log management and its child (lack of adequate logging and regular log review in applications) is to ensure adequate audit log storage.
- ii. The mitigation of data recovery and its child (attacker makes changes to configurations) is through managing data recovery, which includes performing automated backups, protecting recovery data, and testing data recovery.
- iii. The mitigation network infrastructure management and its child (insecure network infrastructure) is to ensure that network infrastructure is up to date.

5.8 Elevation of Privilege

5.8.1 Scenario Description of Elevation of Privilege

Table 5.6 shows the security goal, the source of security threats for the elevation of the privilege component, and how insurance chatbots are vulnerable to this component.

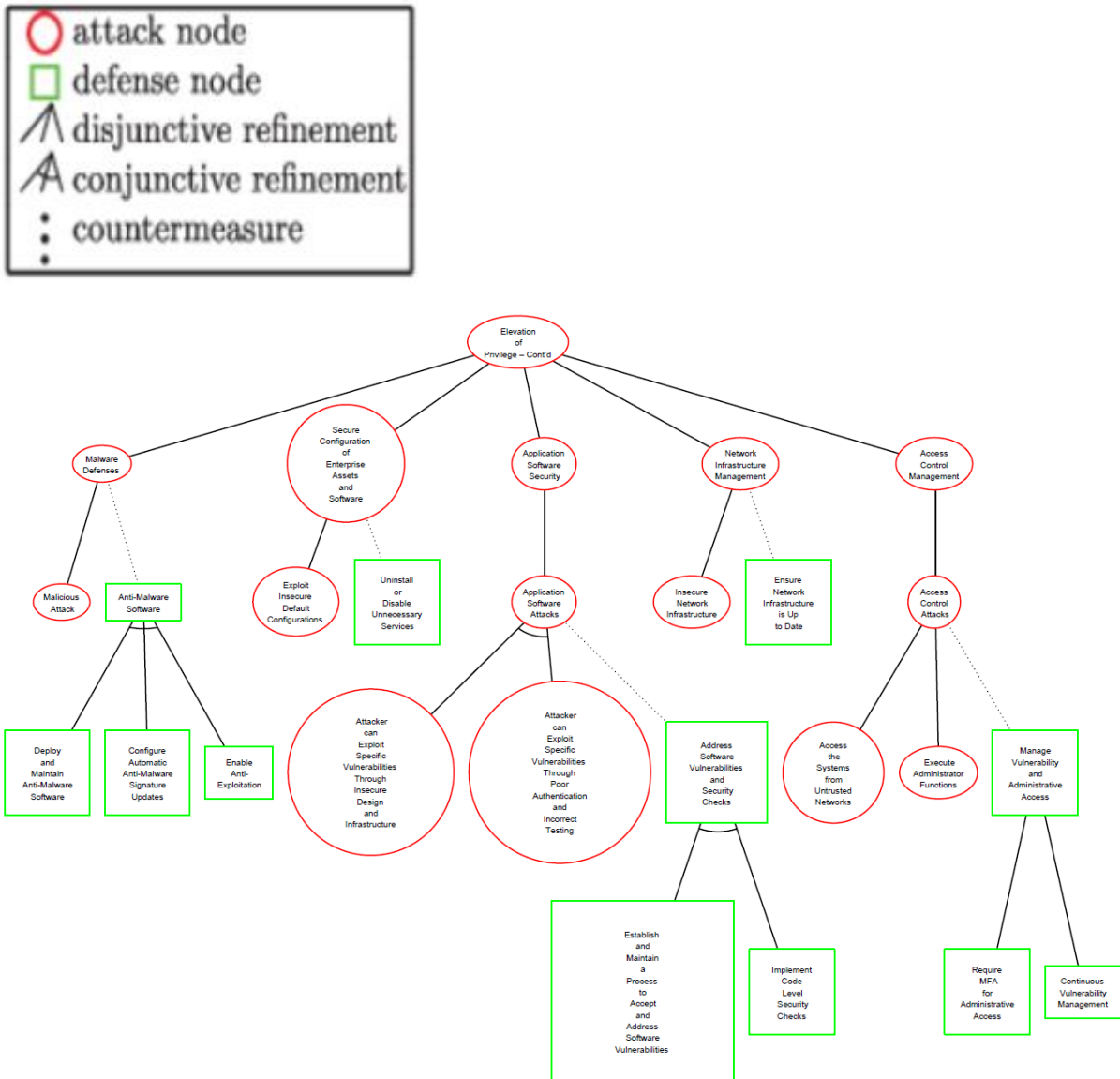
Table 5.6: Elevation of Privilege

	A1	A2	A3	A4	A5
Security Goal					
To provide for the confidentiality, integrity, and availability of the organisation's information assets by ensuring that users who are authenticated cannot obtain privileges for which they are not authorized, access unauthorized information assets, and/or transact in an unauthorised manner.	√	√	√	√	√
Sources of Security threats					
Common Adversarial Threats such as Disgruntled/disaffected employees, Dissatisfied customers, Extremist groups/terrorists, Hackers/hacking groups, Hacktivists, Investigative journalists, Nation-states, Organised criminal groups, Rogue suppliers/vendors/partners, Unscrupulous competitors	√	√	√	√	√
Security Vulnerabilities					
Loopholes in the authorisation mechanisms	√	√	√	√	√
Forms <ul style="list-style-type: none"> • Bypassing authorisation checks • Manipulating existing authorised processes • Privilege escalation • Perform forced browsing/navigation 					
Misconfiguration affecting: <ul style="list-style-type: none"> - End-user systems, Database servers - Web servers and database management systems - Operating Systems - Virtual systems - Networking equipment - Mobile devices 	√	√	√	√	√

Security loopholes	√	√	√	√	√
Forms: <ul style="list-style-type: none"> • Adverse application behaviour/performance • Accessing and getting unauthorised information • Using email and instant messaging • Committing fraud • Unauthorised modification of data, • Steal and disclose enterprise information assets 					
Unauthorised scanning or probing of a company's data applications	√	√	√	√	√
Unauthorised analysis of publicly available information about an enterprise	√	√	√	√	√
Coding bugs or poor design	√	√	√	√	√
Affecting: <ul style="list-style-type: none"> - End-user systems, database servers - Web servers and database management systems - Operating Systems - Virtual systems - Networking equipment - Mobile devices Forms: <ul style="list-style-type: none"> • Buffer overflows • Imperfect validation of input 					
Possible Security Threats					
An attacker can scan for vulnerabilities to exploit the remote software	√	√	√	√	√
Hackers gain unauthorised access	√	√	√	√	√
Proactive attack on newly published vulnerabilities	√	√	√	√	√
An attacker can target unknown vulnerabilities	√	√	√	√	√
An attacker can bypass security controls	√	√	√	√	√
Malicious attack	√	√	√	√	√
An attacker can exploit insecure default configuration	√	√	√	√	√

An attacker can exploit a specific vulnerability through insecure design and infrastructure	√	√	√	√	√
An attacker can exploit a specific vulnerability through poor design and incorrect testing	√	√	√	√	√
Insecure network infrastructure	√	√	√	√	√
An attacker can access the system from an unauthorised network	√	√	√	√	√
An attacker can execute administrator functions	√	√	√	√	√

5.8.2 Threat Model for Elevation of Privilege



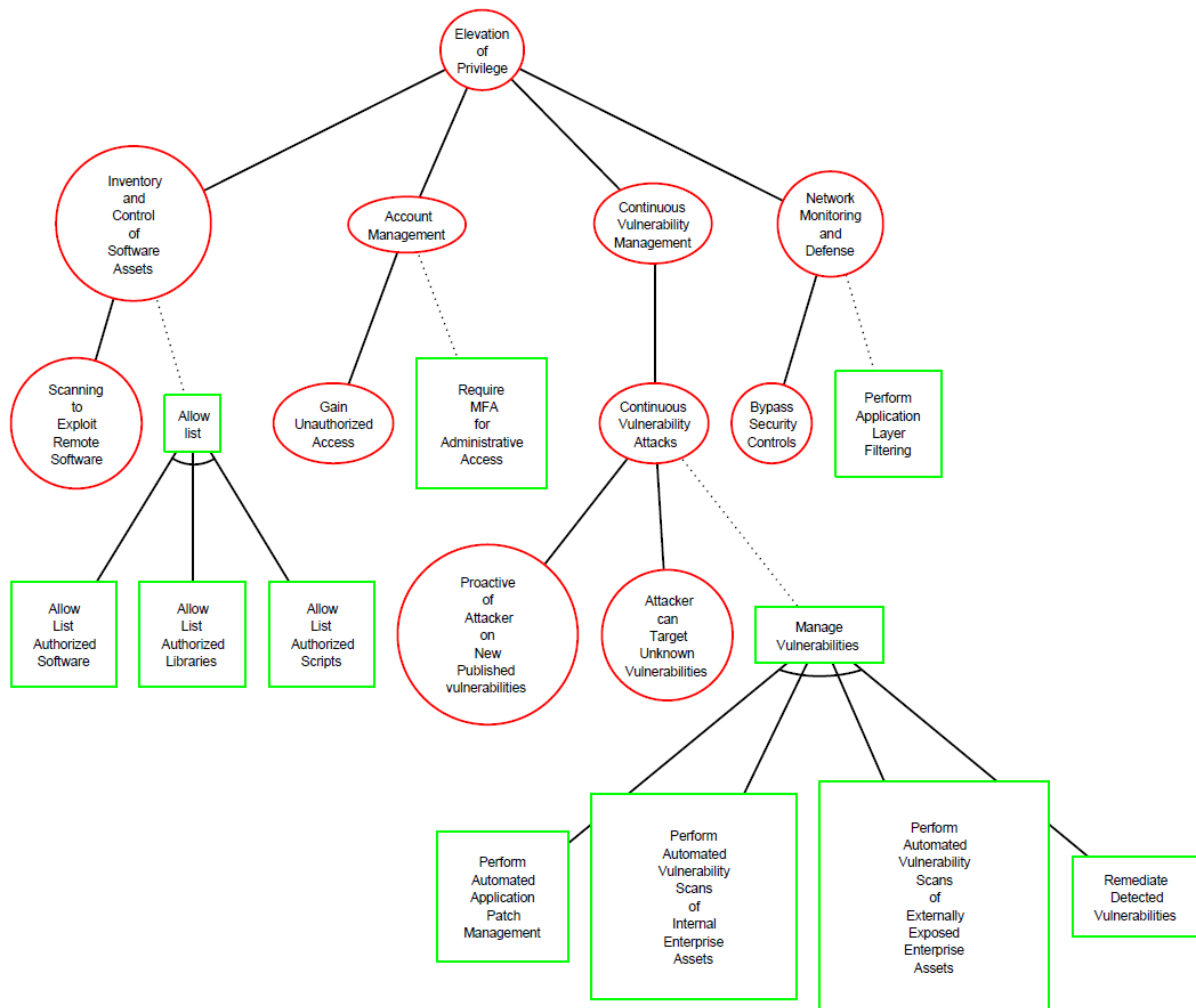


Figure 5.6: Elevation of Privilege Attack and Defence

5.8.3 Description of the Threat Model for Elevation of Privilege

According to Figure 5.6, Elevation of Privilege attacks can be through inventory and control of software assets, account management, continuous vulnerability management, network monitoring and defence, malware defences, secure configuration of enterprise assets and software, application software security, network infrastructure management, and access control management.

- i. The mitigation for Inventory and control of software assets and its child (scanning for vulnerabilities to exploit remote software) is through allowing a list of authorised software, allowing a list of authorised libraries, and allowing a list of authorised scripts.
- ii. The mitigation for account management and its child (gain unauthorised access) is requiring MFA for administrative access.
- iii. The mitigation for continuous vulnerability management and its child: continuous vulnerability attacks which also have two children (proactive attack on newly published vulnerabilities, an attacker can target unknown vulnerabilities) is through managing vulnerabilities, including performing automated application patch management, performing automated application patch

management, performing automated vulnerability scans of internal enterprise assets, performing automated vulnerability scans of externally-exposed enterprise assets, and remediating detected vulnerabilities.

- iv. The mitigation for network monitoring and defence and its child (bypass security controls) is to perform application layer filtering.
- v. The mitigation for malware defences and its child (malicious attack) is through anti-malware software which includes deploying and maintaining anti-malware software, configuring automatic anti-malware signature updates, and enabling anti-exploitation.
- vi. The mitigation for secure configuration on enterprise assets and software and its child (exploit insecure default configurations) is to uninstall or disable unnecessary services.
- vii. The mitigation for application software security management and its child: application software attacks which has also two children (attacker can exploit specific vulnerabilities through insecure design and infrastructure, attacker can exploit specific vulnerabilities through poor authentication and incorrect testing) is through addressing software vulnerabilities and security checks, which include establishing and maintaining a process to accept and address software vulnerabilities, and implementing code level security checks.
- viii. The mitigation network infrastructure management and its child (insecure network infrastructure) to ensure network infrastructure is up to date.
- ix. The mitigation for access control management and its child: access control attacks which also has two children (access the systems from untrusted networks, execute administrator functions) is through managing vulnerability and administrative access, which includes requiring MFA for administrative access, and continuous vulnerability management.

5.9 Chapter Summary

In this chapter, the data was collected from security experts. STRIDE modelling was used as a tool for data collection. The data collected through STRIDE was analysed and the proposed threat model for the study was developed using the AD tool.

CHAPTER SIX

THREAT MODEL EVALUATION

This chapter presents the evaluation of the threat model for data security in insurance chatbots. The evaluation involved presenting the threat model for evaluation and feedback by security experts and chatbot developers. The feedback includes the response to a usability questionnaire and general review comments. The results of the evaluation are presented in this section.

6.1 Profile of the Evaluators

Three security experts and two chatbot developers participated in the threat model evaluation as outlined in the below table. Participant One has been working in the insurance industry since 2000, so going on for 23 years now, and has been involved with Information Security since 2005. Participant 2 has been in the insurance industry since 2005. The participant joined the first insurance industry from 2005 to 2013 and joined the current insurance industry which is the case study in 2013. In both companies, the participant was the Information Security Officer. Participant Three has been working for the insurance industry for 4 years as an information security consultant. Participant Four is the software developer and the case study is the second insurance industry the participant has been working in, with 8 years of experience working in security. Participant 5 is a software developer; the case study is the only insurance company the participant has been working for and has 10 years of experience working in security.

Table 6.1: Profile of Participants in the Evaluation Experiment ~~Thread Evaluation Participants~~

	Level of experience in the insurance industry	Level of experience in security
Participant 1	23 years	18 years
Participant 2	18 years	18 years
Participant 3	4 years	4 years
Participant 4	6 years	8 years
Participant 5	4 years	10 years

6.2 The Evaluation Process

The study adopted the System Usability Scale (SUS) to do the evaluation. SUS is one of the frequently used questionnaires to measure the usability of a system or product. It consists of ten questions. Every odd-numbered question is positively framed, and every even-numbered question is negatively framed (Adrian, 2013; Vlachogianni & Tselios 2022).

The questionnaire of the study was designed using the SUS approach to assess the threat model that has been proposed for data security in chatbots within the insurance industry based on STRIDE – Spoofing (S), Tempering with Data (T), Repudiation (R), Information Disclosure (I), Denial of Service (D), Elevation of Privilege (E).

For each item, the response indicated on the Likert Scale (1-5) below:

1-Strongly Disagree (SD); 2-Disagree (D); 3-Neutral (N); 4-Agree (A); 5-Strongly Agree (SA)

6.3 Evaluation using the System Usability Scale

Table 6.2 shows the structure of questions that were sent to participants to evaluate the threat model.

Table 6.2: SUS Questionnaire for the Threat Model Evaluation

Question Items		SA	D	N	A	SA
		1	2	3	4	5
1	I think I would like to use the proposed threat model					
2	I find the proposed threat model unnecessarily complex					
3	I think the proposed threat model is easy to use					
4	I think I would need the support of security experts to be able to understand and use the proposed threat model					
5	I found the identified threats and suggested mitigations in the model well integrated					
6	I think there is too much inconsistency in this threat model					
7	I think most people will learn to use this threat model very quickly					
8	I find the threat model very cumbersome to use					
9	I feel very confident in using the threat model					
10	I need to learn a lot of things before I use the proposed threat model					

According to Adrian (2013), the SUS score should be interpreted based on the outline shown in Table 6.3

Table 6.3: SUS Score Interpretation

SUS Score	Grade	Adjectival Rating
>80.3	A	Excellent
68-80.3	B	Good
68	C	Ok
51-68	D	Awful
-51	E	Poor

6.4 Evaluation Results

To calculate the SUS score, there are specific steps to be followed (Adrian, 2013, Vlachogianni et al, 2022). The steps below show the evaluation results process from questionnaires 1 to 5. The first step converts all the scales of odd-numbered questions into numbers then again the scales of even-numbered questions into numbers. The second step calculates the sum of odd numbers and assigns them to X then calculate the sum of even numbers and assigns it to Y. The third step calculates the SUS score. To calculate the SUS score, first minus the sum of X by 5 and assign it to XO, then minus 25 by the sum of Y and assign it to YO. Sum XO and YO, multiply the total by 2.5, and then assign it to SUS. The calculation of the SUS score ends up with a number between 1 and 100. After calculating the SUS score for each questionnaire the study calculates the average of all SUS scores from all five questionnaires. With the average total, the study gives the grade and the rating to the developed data security threat model based on the SUS score interpretation table. The table below shows the evaluation results of the developed threat model.

Table 6.4: SUS Score

Participants	Total Score for Odd Questions (X)	Total Score for Even Questions (Y)	Calculation of XO; YO	SUS Score	Total SUS Score
P1	23	10	XO = 23 – 5 = 18 YO = 25 – 10 = 15	18+15 = 33*2.5	82.5
P2	23	10	XO = 23 – 5 = 18 YO = 25 – 10 = 15	18+15 = 33*2.5	82.5
P3	21	25	XO = 21 – 5 = 16 YO = 25 – 14 = 11	16+11 = 27*2.5	67.5
P4	30	25	XO = 30 – 5 = 25 YO = 25 – 13 = 12	25+12 = 37 * 2.5	92.5

P5	17	25	XO = 17-5 = 12 YO = 25-8 = 17	12+17 = 29 * 2.5	72
Mean SUS Score					79.4

The threat model for data security in chatbots obtained an average SUS score (79.4) which can be interpreted as grade B (Good) based on the SUS interpretation table.

6.5 Security Experts' Feedback

The security experts were also asked to give general comments and observations (in the form of a narrative) on the quality of the threat model. Security experts gave positive feedback that the developed threat model is a valuable contribution to the field of information security. They commented that it is good to see security frameworks like the Attack Defence tool and STRIDE being used in the study. They observed that by using these tools, organisations can better understand the vulnerabilities and risks associated with the chatbot and take steps to address them. This can help to improve overall security and protect against potential attacks, and also proactively identify and mitigate potential threats

The security experts also emphasised that the developed threat model is high-quality, thorough, and comprehensive, accurately identifying, and prioritising potential threats to an organisation's chatbots, the threat model is also based on a solid understanding of the system being modelled, including its architecture, design, and operation, as well as the potential adversaries and their motivations and capabilities.

6.6 Threat to Validity

In this study, validity means a true and accurate representation of information. The validity of the developed threat model is based on discussing the validity of the experiment, including conclusion validity, internal validity, construct validity, and external validity (Wohlin et al., 2012; Emebo, Daramola, and Ayo, 2017).

Conclusion Validity- To draw a reasonable conclusion based on an analysis of the data, all participants were provided with a data security threat model and questionnaires. All the questionnaires were designed using System Usability Scale so that the rating of the threat model can be accurate. The structure and instructions in all questionnaires were the same. Although the case study organisation is in a multicultural environment, the main language of the organisation is English; so, all questionnaires and instructions were written in English.

Internal Validity- To have the ability to conclude the causal relationship from the data, five participants with a good experience in data security were provided with a threat model and questionnaires. The questionnaire included detailed instructions on the questions that should be answered. Also, the

participants were instructed to contact the researcher should they require more information. All participants have a minimum level of education in the area of data security. Considering the response of the security experts, the conclusion can be drawn that the threat model will contribute positively to data security in chatbots in the South African insurance industry.

Construct Validity- Refers to the adequacy of the operational definition of variables. To ensure the validity of the threat model rating, participants were provided with the same questionnaires. The participants followed the same instruction as the guide in rating the threat model. The results obtained from the survey for a rating of the threat model depend on one variable which removes the non-methods bias effect.

External Validity- To generalise the results to other populations and settings. All participants are from the same case study, which is a single case study. Two participants are chatbot developers and three are chatbot experts. Considering that all participants have the same understanding of data security for chatbots since they are from the same case study, the concern is that there may be different results if the evaluation was done with a group of more than five participants and from multiple case studies with more diverse data security background. On another hand, The concern does not mean the threat model cannot be used, but it indicates that for future research, it will be good that the focus would be on multiple case studies rather than a single case study.

To summarise, there are no major concerns about the validity of the developed threat model. Also, the fact that no other study has developed a threat model for data security in small and medium-sized insurance organisations makes the developed threat model a great contribution to data security in chatbots in South African insurance organisations.

6.7 Chapter Summary

This chapter presents the evaluation of the validity of the developed threat model. SUS Questionnaires for threat evaluation were handed out to the security expert. The security experts were able to do the evaluation and gave general comments and observations (in the form of a narrative) on the quality of the threat model. SUS score was calculated from the evaluation. Based on the score (79.4) the rated grade is B which is interpreted as a good rating in the interpretation of the SUS score. The validity of the developed threat model was also discussed.

CHAPTER SEVEN

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

The research summary, research conclusion, contribution, future research, and recommendations of the study are discussed in this chapter. The summary section gives the overview of what was covered in this study starting from objective one to objective five. The conclusion of the study gives the judgment or decision reached by the study. The recommendation section indicates what needs to be covered by future research on the subject. The contribution section explains how the study contributed to the body of knowledge.

7.1 Research Summary

This research aimed to develop a threat model for the security of data in chatbots used in insurance organisations. The study was divided into seven chapters since each chapter contributes differently to the study.

Chapter 1 discussed the introduction, background, research problem, aim, objectives, and research questions; delineation, and significance of the study. Chapter 2 reviewed the literature relevant to this study. It started by discussing the chatbot applications, chatbot security, and the insurance industry in South Africa and then discusses thread modelling, which is described as a technique used in the identification of security threats. Chapter 3 presented the research methodology by discussing the research process used in this study. Chapter 4 presented the findings of the analysis of data collected from the case study. Chapter 5 presented the development of the proposed threat model for data security in chatbots. Chapter 6 presented the threat model evaluation process by security experts. Chapter 7 presents the summary, recommendation, and conclusion of the study.

How the objectives of the study were achieved is elaborated below:

i. **Objective 1: To identify the potential use cases of chatbots for CRM in a South African insurance organisation.**

Following the research design, this objective was achieved by conducting interviews in the selected case study and reviewing existing literature. Interviews were conducted involving the stakeholders in an insurance organisation.

During the data collection, ten participants were interviewed and the participants were purposively selected. The roles and officers that were purposively selected to participate in the research are as follows: security experts, chatbot developers, chatbot testers, chatbot users, and chatbot managers. Everyone participating in this research received a consent letter which is a written communication from the Cape Peninsula University of Technology that explains that the researcher has the approval for collecting information to carry out the study. Also in the consent, participants were informed about the nature, actions, risks, benefits, etc., of the research in a way that is not technical and was easy to understand.

The data collected from the participants were analysed and grouped into themes. From the data, chatbot use cases in the insurance industry were identified as the outcome of this objective. The chatbot use cases are as follows: Claims iAssist bot, Human Resource iAssist bot, Personal Lines iAssist bot, Commercial Lines iAssist bot, and Policy WhatsApp bot. The data flow in each chatbot was presented.

ii. Objective 2: To identify the challenges of securing data in a chatbot in a South African insurance organisation.

This objective was achieved by conducting interviews in the case study and reviewing existing literature. Interviews were conducted involving the stakeholders in an insurance organisation.

The data collected from the participants were analysed and grouped into themes. From the collected and analysed data, data security challenges in chatbots in the insurance industry were identified as the outcome of this objective. The current security challenge in chatbots is that it can be possible for an attacker to perform the following security attacks: Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege.

- Spoofing- An attacker can illegally access and use another user's credentials. Impersonating something or someone else.
- Tampering- An attacker can maliciously change/modify data.
- Repudiation- An attacker can perform an illegal operation in a system
- Information Disclosure- An attacker can steal confidential information
- Denial of Service- An attacker can deny access to valid users.
- Elevation of Privilege- An attacker can gain privileged access.

iii. Objective 3: To determine the security goals, threats, and vulnerabilities associated with the use of chatbots in a South African insurance organisation.

This objective was achieved by using STRIDE modelling as threat elicitation to identify security goals, possible threats, and vulnerabilities of each chatbot in the organisation regarding the different components of STRIDE based on the perspectives of the expert. The experts were presented with a document template to capture their individual analysis of security goals, threats, and vulnerabilities, and thereafter areas of consensus were noted, while areas of differences were resolved in joint meetings of the experts. This led to the final documentation of the security goals, threats, and vulnerabilities associated with the use of chatbots in the insurance organisation.

iv. Objective 4: To develop a threat model for the security and privacy of data in chatbots for a South African insurance organisation.

This objective was achieved by developing a threat model based on the data collected from STRIDE modelling. The threat model diagrams examined whether each chatbot and its related asset data is vulnerable to any of the security threats in STRIDE and it also showed the actions of an attacker trying to compromise the system and possible counteractions of a defender trying to protect the system. It also helped with the qualitative analysis of security, using attack–defence trees. The diagrams were developed by using an Attack Defence tool called ADTool.

v. Objective 5: To evaluate the threat model for the security and privacy of data in the chatbots for a South African insurance organisation.

To achieve this objective with the research design approach, the evaluation of the study adopted the System Usability Scale (SUS), which is a tool to evaluate a system or product. The questionnaires were designed using the SUS approach to assess the threat model proposed for data security in chatbots within the insurance industry based on STRIDE. For each item, the response indicated on the Likert Scale (1-5) below: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree. The security experts evaluated the threat model and gave quality ratings of the threat model. They also provided qualitative feedback in the form of general comments.

7.2 Limitations of the Study

The limitation of the study is discussed below as access and methodological limitations.

Access limitations: The permission for data collection in the case study was granted without any difficulties. Using a single case study was also an advantage since permission for data collection was only requested from one organisation. The only difficulties were during the interview process where it was very difficult to get hold of participants in managerial positions because of their busy schedules, which delayed the data collection activities in this study.

Methodological limitations: During the data collection, the study had to change from qualitative to mixed method methodology because the data was both textual and numerical. A single case study was a research approach for this study since the time of the research was limited. A different approach, which is a multiple case study, would be more suitable if there had been time and resources for this research since different results would come from data collected from participants with more diverse chatbots and data security backgrounds.

7.3 Contributions of the Study

The contribution of the study is conceptualised in terms of the following: Theoretical contribution and Practical contribution.

7.3.1 Theoretical Contribution

This study explores the issue of data security in chatbots in the insurance industry which has not received much attention so far in the literature. The fact that it is focused on the South African context also adds to an existing body of knowledge on chatbot security.

7.3.2 Practical Contribution

The study developed a threat model for data security in chatbots that will be used in practice by the insurance industry in South Africa. The study contributes to both organisations and academics. As organisations can use the study as a security strategy when implanting chatbots and for academics, it contributes to the body of knowledge.

7.3 Future Research and Recommendation

In this study, a threat model for data security in chatbots was developed and evaluated based on security experts' perspectives. The evaluation was not based on practical usage or deployment of the threat model. Future research can focus on impact assessment and in-use evaluation of the threat model as the security expert will have more insight into the developed threat model when doing the evaluation.

The threat model that was developed was based on STRIDE modelling. Future research of the study could use other data security models such as Abuser stories (Peeters, 2005; Crothers et al., 2023), STRIDE average model (Jesan, 2008; Zaeni et al., 2023), Attack trees (Satapathy, 2014; Ebrahimi et al., 2022), Fuzzy Logic (Sodiya, Onashoga, Oladunjoye, 2007; Batool et al., 2022), SDL Threat Modeling tool (Shostack, 2008; Santa, 2023), T-map (Lodderstedt, Basin, & Doser, 2002; Hu et al., 2022), and CORAS (Hussain, Kamal, Ahmad, Rasool & Iqbal, 2014; Heisel et al., 2023) which are also some of several Microsoft Threat modelling methodologies and techniques for identifying security threats.

In future work, the Cyber Security Framework (CSF), which was created by The National Institute of Technology (NIST, 2014) as a voluntary framework to be used by organisations as a strategy for preventing, detecting, and responding to cyberattacks can be used to create a threat model for chatbot data security. The CSF also has a notation that can be used to present threats and mitigation instead of using the AD Tool that was used in the study.

REFERENCES

- AbdulRaheem, M., Awotunde, J.B., Chakraborty, C., Adeniyi, E.A., Oladipo, I.D. and Bhoi, A.K., 2023. Security and privacy concerns in smart healthcare system. In *Implementation of Smart Healthcare Systems using AI, IoT, and Blockchain* (pp. 243-273). Academic Press.
- Abend, G. 2008. The Meaning of Theory. *Sociological Theory*, 26(2):173 - 199
- Alavudeen, R. & Rosa, K. D. 2015. Growing Role of Bancassurance in the Banking Sector. *Bonfring International Journal of Industrial Engineering and Management Science*, 5(2):10-16
- Alwahaby, H., Cukurova, M., Papamitsiou, Z. and Giannakos, M., 2022. The evidence of impact and ethical considerations of Multimodal Learning Analytics: A Systematic Literature Review. *The Multimodal Learning Analytics Handbook*, pp.289-325.
- Arias, J., Petrucci, L., Masko, L., Penczek, W. & Sidoruk, T. 2022. Minimal Schedule with Minimal Number of Agents in Attack-Defence Trees.
- Naz, N., Gulab, F. and Aslam, M., 2022. Development of qualitative semi-structured interview guide for case study research. *Competitive Social Science Research Journal*, 3(2), pp.42-52.
- Balasubramanian, R., Libarikian, A. & McElhaney, D. 2018. Insurance 2030—The impact of AI on the future of insurance. <https://www.mckinsey.com/industries/financial-services/our-insights/insurance-2030-the-impact-of-ai-on-the-future-of-insurance> [30 Aprile 2018]
- Bangera, S. and Bhat, S., 2023. A Systematic Study of Application of Cognitive Intelligence in Mphasis—a Case Study. *International Journal of Case Studies in Business, IT and Education (IJCSBE)*, 7(2), pp.360-380.
- Barriball, L. K. & While, A. 1994. Collecting Data Using a Semi-Structured Interview: A Discussion Paper. *Journal of Advanced Nursing*, 19, 328-335. <http://dx.doi.org/10.1111/j.1365-2648.1994.tb01088.x>
- Batool, A., Hussain, M. and Abidi, S.M.R., 2022. Intelligent Cloud Security Issues Detection Using Mamdani Fuzzy Logic. *International Journal of Computational and Innovative Sciences*, 1(3), pp.33-51.
- Bhuiyan, M.S.I., Razzak, A., Ferdous, M. S., Chowdhury, M. J. M., Hoque, M. A., Tarkoma, S (2020) BONIK: A Blockchain-Empowered Chatbot for Financial Transactions. *IEEE 19th International Conference on Trust, Security, and Privacy in Computing and Communications* 1079–1088 <https://doi.org/10.1109/trustcom50675.2020.00143>
- Bhuiyan, M.S.I., Razzak, A., Ferdous, M. S., et al. 2020. BONIK: A Blockchain-Empowered Chatbot for Financial Transactions. *IEEE 19th International Conference on Trust, Security, and Privacy in Computing and Communications* 1079–1088. <https://doi.org/10.1109/trustcom50675.2020.00143>

- Bozic, J. & Wotawa, F. 2018. Planning-based Security Testing for Chatbots. 30th IFIP International Conference on Testing Software and Systems, ICTSS 2018. Spain, 1 Oct 2018 - 3 Oct 2018: 33-38 [1 Jan 2018]
- Bryman, A. & Bell, E. 2015. Business Research Methods. 4th edition. Cambridge, United Kingdom; New York, NY, United States of America: Oxford University Press.
- Cavanillas, J. M., Wahlster, W. & Curry, E. 2016. New Horizons for a Data-Driven Economy. A Roadmap for Usage and Exploitation of Big Data in Europe: Springer Nature
- Crothers, E., Japkowicz, N. and Viktor, H.L., 2023. Machine-generated Text: A Comprehensive Survey of Threat Models and Detection Methods. *IEEE Access*.
- Creswell, J.W. 2013. Research design: Qualitative, quantitative, and mixed methods approach. Thousand Oaks, California: SAGE Publications.
- Ebrahimi, M., Striessnig, C., Triginer, J.C. and Schmittner, C., 2022. Identification and verification of attack-tree threat models in connected vehicles. *arXiv preprint arXiv:2212.14435*.
- Emebo, O., Daramola, O. & Ayo, C. 2017. A Survey on Implicit Requirements Management Practices in Small and Medium-Sized Enterprises.
- Esposito, S., Santis, A. D., Tortora, G., Chang, H. & Choo, K. R. 2018. Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy? DEPARTMENT: Cloud and the Law
- Cummins, J. D., Tennyson, S. & Weiss, M. A. 1998. Efficiency, Scale Economies, and Consolidation in the U.S. Life Insurance Industry. Center for Financial Institutions Working Papers. United States: 98-08.
- Da Silva, M., Puys, M., Thevenon, P.H., Mocanu, S. and Nkawa, N., 2023, August. Automated ICS template for STRIDE Microsoft Threat Modeling Tool. In *Proceedings of the 18th International Conference on Availability, Reliability and Security* (pp. 1-7).
- Ditizio, A. A. & Smith, A. D. 2017. Transformation of CRM and Supply Chain Management Techniques in a New Venture. <https://www.igi-global.com/chapter/transformation-of-crm-and-supply-chain-management-techniques-in-a-new-venture/166517>
- Dusitnanond, A. 2007. Developing a Method of Teaching Architectural Project Design: A Case Study of Third Year Studio Project, Faculty of Architecture, Sriburapha University, Thailand. <http://vuir.vu.edu.au/1571/>
- Edu, J., Mulligan, C., Pierazzi, F., Polakis, J., Suarez-Tangil, G. and Such, J., 2022, October. Exploring the security and privacy risks of chatbots in messaging services. In *Proceedings of the 22nd ACM internet measurement conference* (pp. 581-588).
- Etikan, L. 2016. Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1):1.

Fleischmann, M. & Ivens, B. S. 2019. Exploring the Role of Trust in Blockchain Adoption: An Inductive Approach. Proceedings of the 52nd Hawaii International Conference on System Sciences. 08 Jan 2019. <http://hdl.handle.net/10125/60120>

Følstad, A., Nordheim, C.B. and Bjørkli, C.A., 2018. What makes users trust a chatbot for customer service? An exploratory interview study. In *Internet Science: 5th International Conference, INSCI 2018*, St. Petersburg, Russia, October 24–26, 2018, Proceedings 5 (pp. 194-208). Springer International Publishing.

Garcia, A. 2013. UX Research | Standardized Usability Questionnaire.

<https://chaione.com/blog/category/user-research/>

Gill, P., Stewart, K., Treasure, E. & Chadwick, B. 2008. Methods of data collection in qualitative research: Interviews and focus groups. *British dental journal official journal of the British Dental Association: BDJ online*, 204(6):291-5.

Gondaliya, K., Butakov, S. & Zavorsky, P. 2020. SLA AS A MECHANISM TO MANAGE RISKS RELATED TO CHATBOT SERVICES

Goddard, W. & Melville, S. 2004 *Research Methodology: An Introduction*. Juta Academic.

Grand View Research: Chatbot Market Size To Reach \$1.25 Billion By 2025 (2017). <https://www.grandviewresearch.com/press-release/global-chatbot-market>

Guillem, M. A. L. 2022. INSURANCE BROKERS' BEHAVIOUR: THE EFFECT OF POLICY COLLECTION ON MANAGEMENT DECISIONS. Catholic University of Valencia "San Vicente Mártir", Spain

Gustafsson, J. 2017. Single case studies vs. multiple case studies: A comparative study.

Harkous, H., Shin, K. G., Fawaz, K. & Aberer, K. 2016. PriBots: Conversational Privacy with Chatbots

IBM, 2017. <https://www.ibm.com/blogs/watson/2017/10/how-chatbots-reduce-customer-service-costs-by-30-percent/>

Hasal, M., Nowakova, J. Saghair, K. A., Abdulla, H., Snasel, V. & Ogiela, L. 2021. Chatbots: Security, privacy, data protection, and social aspects

Heisel, M. and Wagner, M., 2023. Pattern-Based Risk Identification for Model-Based Risk Management. In *Applicable Formal Methods for Safe Industrial Products: Essays Dedicated to Jan Peleska on the Occasion of His 65th Birthday* (pp. 114-129). Cham: Springer Nature Switzerland.

Hu, S., Zhou, Z., Zhang, Y., Zhang, L.Y., Zheng, Y., He, Y. and Jin, H., 2022, October. Badhash: Invisible backdoor attacks against deep hashing with clean label. In *Proceedings of the 30th ACM International Conference on Multimedia* (pp. 678-686).

Huang, Y.S.S. and Dootson, P., 2022. Chatbots and service failure: When does it lead to customer aggression. *Journal of Retailing and Consumer Services*, 68, p.103044.

Isinkaye, F.O., AbiodunBabs, I.G. and Paul, M.T., 2022. Development of a Mobile-Based Hostel Location and Recommendation Chatbot System. *IJ Information Technology and Computer Science*, pp.23-33.

Jamin, J., Arifin, N.A.M., Mokhtar, S.A., Rosli, N.N.I.N., Shukry, A.I.M. 2019. Privacy Concern Of Personal Information In The ICT Usage, Internet, and Social Media Perspective. *Malaysian E Commerce Journal (MECJ)*, 3(2): 15-17.

Jesan, P., 2008. "Threat modeling web applications using STRIDE average model." *Computer Security Conference*.

Juniper Research, 2018. Chatbots to Deliver \$11bn in Annual Cost Savings for Retail, Banking & Healthcare Sectors by,2023,03-July-2018.[Online].Available:<https://www.businesswire.com/news/home/20180703005029/en/Juniper-Research-Chatbots-Deliver-11bn-AnnualCost>

Khan, R., 2017. Standardized architecture for conversational agents aka chatbots. *International Journal of Computer Trends and Technology*, 50(2), 114-121.

Koetter, F., Blohm, M., Kochanowski, M., Goetzer, J., Graziotin, D. & Wagner, S. 2018. Motivations, Classification, and Model Trial of Conversational Agents for Insurance Companies.

Kumar, R. 2019. *Research Methodology: A Step-by-Step Guide for Beginners*.

Lai, S.T., Leu, F.Y. and Lin, J.W., 2019. A banking chatbot security control procedure for protecting user data security and privacy. In *Advances on Broadband and Wireless Computing, Communication and Applications: Proceedings of the 13th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA-2018)* (pp. 561-571). Springer International Publishing.

Lechner, N.H., Strahonja, V. and Stapić, Z., 2022. Threat modeling methods in the medical device industry: An integrative literature review. In *Central European Conference on Information and Intelligent Systems* (pp. 479-488). Faculty of Organization and Informatics Varazdin.

Ledro, C., Nosella, A. & Vinelli, A. 2022. Artificial intelligence in customer relationship management: literature review and future research directions. *Journal of Business & Industrial Marketing* 37/13 (2022) 48–63 Emerald Publishing Limited, ISSN 0885-8624, DOI 10.1108/JBIM-07-2021-0332

Lee, P., Bubeck, S. & Petro, J. 2023. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine.

Lin, C., Huang, A. Y. Q., Stephen, J. H. Y. 2023. A Review of AI-Driven Conversational Chatbots Implementation Methodologies and Challenges (1999–2022). *Sustainability* 2023, 15, 4012. <https://doi.org/10.3390/su15054012>

- Lim, W. M., Kumar, S. & Ali, S. 2022. Advancing knowledge through literature reviews: 'what', 'why', and 'how to contribute', *The Service Industries Journal*, 42:7-8, 481-513, DOI: 10.1080/02642069.2022.2047941. <https://doi.org/10.1080/02642069.2022.2047941>
- Liu, Y., Peng, J. & Yu, Z. 2019. Big Data Platform Architecture Under The Background of Financial Technology.
- Lodderstedt, T. Basin, D. & Doser, J. 2002. SecureUML: A UML-based modeling language for model-driven security. «UML» 2002—The Unified Modeling Language, 426–441.
- Namasudra, S., Devi, D., Kadry, S. Sundarasekar, R & Shanthini, A. 2020. Towards DNA-based data security in the cloud computing environment. *Computer Communications*. 151 (2020) 539–547
- Ng, M., Coopamootoo, K.P., Toreini, E., Aitken, M., Elliot, K. and van Moorsel, A., 2020, September. Simulating the effects of social presence on trust, privacy concerns & usage intentions in automated bots for finance. In 2020 IEEE European symposium on security and privacy workshops (EuroS&PW) (pp. 190-199). IEEE.
- Magano, K. D. & de Beer. K. T. 2021. Investigating change fatigue, burnout, work engagement, organisational commitment and turnover intention in the South African insurance industry. Mini-dissertation accepted in partial fulfilment of the requirements for the degree Master of Arts in Industrial and Organisational Psychology at the North-West University. orcid.org/0000-0002-9005-4621
- Mas, N. 2011. 4 major types of qualitative research, s.l.: s.n.
- Mott, N. 2018. Ticketmaster Blames Malware-Plagued Chatbot for Data Breach, <https://www.tomshardware.com/news/ticketmaster-data-breach-uk-international,37383.html>
- Murugesan, S. 2019. *The Cybersecurity Renaissance: Security Threats, Risks, and Safeguards*.
- Muzari, T., Shava, G.N. and Shonhiwa, S., 2022. Qualitative research paradigm, a key research design for educational researchers, processes and procedures: A theoretical overview. *Indiana Journal of Humanities and Social Sciences*, 3(1), pp.14-20.
- Ofori-Boateng, K., Ohemeng, W., Ahawaadong Boro, E. and Kwame Agyapong, E., 2022. Efficiency, market structure and performance of the insurance industry in an emerging economy. *Cogent Economics & Finance*, 10(1), p.2068784.
- Ondrisek, B., 2016. Why You Shouldn't Talk to Your Chatbot about Everything. <http://venturebeat.com/2016/11/17/why-you-shouldnt-talk-to-your-chatbot-about-everything/>
- Osmanbegovic, E. & Zahirović, S., 2013. Perception of Information Security of Management of Banking and Insurance Companies in Countries of Western Balkans. Article in *Research in Applied Economics*, 5(2). <https://www.researchgate.net/publication/314507777>
- Peeters, Johan. (2005) "Agile security requirements engineering." Symposium on Requirements Engineering for Information Security.

- Poirier, F., Mansouri, K. and Kaiss, W., 2023, July. Chatbot design to help learners self-regulate their learning in online learning environments. In *The 23rd IEEE International Conference on Advanced Learning Technologies (ICALT 2023)*.
- Pole, C.J., & Lampard, R., 2013. Practical Social Investigation Quantitative and Qualitative Methods in Social Research.
- Ranjit, K., 2019. Research Methodology.
- Raikwar, M., Mazumdar, S., Ruj, S. Gupta, S. S., Chattopadhyay, A & Lam, K. 2019. A Blockchain Framework for Insurance Processes. 2018 9th IFIP International Conference on New Technologies, Mobility, and Security (NTMS). 26-28 Feb. 2018. <https://ieeexplore.ieee.org/abstract/document/8328731>
- Rehman, A. A. & Alharthi, K. 2016. An introduction to research paradigms.
- Saeed, S., Ahmed, M.O. & Malik, U. 2017. Role of Information Communication Technology (ICT) in the 21st Century.
- Sakshi, V.M.V.S. and Sharma, V.K., 2023. Investigating The Performance Of Messenger App Security For WhatsApp, Facebook And Instagram Among Indian Users.
- Santa Rosa, A., 2023. *Validating a Threat Model for Smart Home Gateways* (Doctoral dissertation, Politecnico di Torino).
- Shobana, M., Reddy, K. N. K. & Sai, M. R. V. 2023. HEALTH CARE ASSISTANCE CHAT BOT USING NATURAL LANGUAGE PROCESSING. Industrial Engineering Journal ISSN: 0970-2555 Volume : 52, Issue 5, No. 2, May : 2023.
- Shabbir, J. & Anwer, T., 2018. Artificial Intelligence and its Role in Near Future.
- Shostack, A. 2008., Experiences Threat Modeling at Microsoft. In Modeling Security Workshop. Dept. of Computing, Lancaster University, UK.
- Satapathy, s. R. 2014., Threat Modeling in Web Applications.
- Sodiya, S. A., Onashoga, S.A. Oladunjoye, B. A. 2007. Threat Modeling Using Fuzzy Logic Paradigm. Information and Beyond: Part I, 4, 53.
- Swanson & Richard A., 2013. Theory Building in Applied Disciplines.
- Taherdoost, H., 2022. What are different research approaches? Comprehensive Review of Qualitative, quantitative, and mixed method research, their applications, types, and limitations. *Journal of Management Science & Engineering Research*, 5(1), pp.53-63.
- Tang, M. & Niu, Y. 2023. Importance-Performance Analysis of Online Insurance: Communication and Networking. Department of Insurance, Chaoyang University of Technology, Taichung 413310, Taiwan.

Presented at the 3rd IEEE International Conference on Electronic Communications, Internet of Things and Big Data Conference 2023, Taichung, Taiwan, 14–16 April 2023.

Tedesco, S., Barton, J. & O'Flynn, B., 2015. A Review of Activity Trackers for Senior Citizens: Research Perspectives, Commercial Landscape and the Role of the Insurance Industry. *Ensors* **2017**, 17(6), 1277; <https://doi.org/10.3390/s17061277>

Vakilinia, I., Tosh, D. & Sengupta, S., 2017. 3-Way game model for privacy-preserving cybersecurity information exchange framework.

V. Hristidis, 2018., Chatbot Technologies and Challenges, First International Conference on Artificial Intelligence for Industries (AI4I), vol. doi: 10.1109/AI4I.2018.8665692, pp. 126-126, 2018.

Varpio, L., Paradis, E., Uijtdehaage, S. and Young, M., 2020. The distinctions between theory, theoretical framework, and conceptual framework. *Academic Medicine*, 95(7), pp.989-994.

Vlachogianni, P. and Tselios, N., 2022. Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. *Journal of Research on Technology in Education*, 54(3), pp.392-409.

Cardona, D. R., Werth, O., Schönborn, S. & Breitner, M. H., 2019. A Mixed-Methods Analysis of the Adoption and Diffusion of Chatbot Technology in the German Insurance Sector. Proceedings of the 25th Americas Conference on Information Systems (AMCIS). Cancun, Mexico, January 2019.

Wenjun, X. & Lagerström, R., 2019. Threat modeling – A systematic literature review.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M.C., Regnell, B. and Wesslén, A., 2012. Experimentation in Software Engineering: An Introduction (Vol. 6). Springer Science & Business Media.

Wube, H.D., Esubalew, S.Z., Weldesellasie, F.F. and Debelee, T.G., 2022. Text-based chatbot in financial sector: a systematic literature review. *Data Sci. Financ. Econ*, 2(3), pp.232-259.

Yan, M., Castro, P., Cheng, P., & Ishakian, V., 2016. Building a Chatbot with Serverless Computing.

Ye, W. and Li, Q., 2020, November. Chatbot security and privacy in the age of personal assistants. In 2020 IEEE/ACM Symposium on Edge Computing (SEC) (pp. 388-393). IEEE.

Yin, R. K., 2018. Case study research and applications: Design and methods (6th ed.). Thousand Oaks, CA: Sage Publications.

Zaeni, I.A.E., Lestari, D., Handayani, A.N. and Osman, M.K., 2023. Development of Stride Detection System for Helping Stroke Walking Training. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 5(3), pp.159-167.

Zhang, Z. Li, B. & Liu, L. 2023. The impact of AI-based conversational agent on the firms' operational performance: Empirical evidence from a call center, *Applied Artificial Intelligence*, 37:1, 2157592, DOI: 10.1080/08839514.2022.2157592. <https://doi.org/10.1080/08839514.2022.2157592>

Zorn, T.E., Flanagan, A. J. & Shoham, M. D. 2011. Institutional and Noninstitutional

APPENDICES

APPENDIX A: OBJECTIVE ONE INTERVIEW PROCESS

Below is the structure of interview questions that were asked the participants of the study to achieve objectives one and two of the study.

Section A

1. What purposes does the chatbot fulfill in your organisation?
2. In what ways are chatbots used in your organisation?
3. List the specific processes/operations where chatbots are used in your organisation.
4. What kind of support and maintenance is available for chatbots in your organisation?
5. What kind of data is stored in chatbots in your organisation?
6. Who are the chatbot users?

Section B

1. Where are the chatbots used in your organisation hosted?
2. How do the chatbots in your organisation integrate with social media platforms?
3. How does the chatbot platform in your organisation store data after transactions?
4. What happens to used data inside chatbots?
5. What features does your chatbot have? e.g. speech recognition, text-based, or speech to text
6. What kind of security measures does your chatbot have to prevent identity theft?
7. What kind of security does the chatbot have to ensure data privacy?
8. What kind of security measures does your chatbot have to ensure data integrity?
9. What kind of security measures does your chatbot have to prevent unauthorised access?
10. What kind of security measure does your chatbot have for user authentication?
11. What are the security vulnerabilities that you have found with your chatbot?

APPENDIX B: Threat Model Evaluation Questionnaire

Below is the structure of interview questions that were asked by the participants of the study to rate the developed threat model for data security in chatbots for South African insurance organisations.

Question Items	
1	I think I would like to use the proposed threat model
2	I find the proposed threat model unnecessarily complex
3	I think the proposed threat model is easy to use
4	I think I would need the support of security experts to be able to understand and use the proposed threat model
5	I found the identified threats and suggested mitigations in the model well integrated
6	I think there is too much inconsistency in this threat model
7	I think most people will learn to use this threat model very quickly
8	I find the threat model very cumbersome to use
9	I feel very confident in using the threat model
10	I need to learn a lot of things before I use the proposed threat model

APPENDIX C: ETHICAL REQUIREMENTS

Appendix C aimed to fulfill the ethical standards required by the Cape Peninsula University of Technology research code of ethics. Appendix B contains (1) the faculty of Informatics and Design Ethical approval; (2) an introductory letter for the collection of research data from the university and the supervisor; (3) a Request to conduct research and interview participation consent letter approval from the case study of the research; (4) the consent letter that was handed out by the researcher to the participants before the interview process began.

Office of the Research Ethics Committee
Faculty of Informatics and Design
Room 2.09
80 Roeland Street
Cape Town
Tel: 021-469 1012
Email: ndedem@cput.ac.za
Secretary: Mziyanda Ndede

05 May 2021

Zilungile Bokolo
c/o Department of Information Technology
CPUT

Reference no: 215296273/2021/12

Project title: Data Security in Chatbots for the Insurance Industry: A case study of a South African Insurance Company.

Approval period: 05 May 2021 – 31 December 2022

This is to certify that the Faculty of Informatics and Design Research Ethics Committee of the Cape Peninsula University of Technology approved the methodology and ethics of Zilungile Bokolo (215296273) for the MTech Information Technology.

Any amendments, extension or other modifications to the protocol must be submitted to the Faculty Research Ethics Committee for approval.

The Committee must be informed of any serious adverse event and/or termination of the study.



A/Prof I van Zyl
Chair: Research Ethics Committee
Faculty of Informatics and Design
Cape Peninsula University of Technology
vanzyliz@cput.ac.za



head office: 1 sportica crescent tyger valley, belville 7530 - po box 3881 tyger valley 7536
t +27 (0)21 915 7000 • www.santam.co.za • enquiries@santam.co.za

11 May 2021

Prof. Justine Olawande Daramola
Research Supervisor
Department of Informatics and Design
Faculty of Information Technology
Cape Peninsula University of Technology
E-mail : Daramolaj@cput.ac.za

Dear Prof. Daramola

Request to conduct research and interview participation consent letter

I, Kevin Wright, in my capacity as Chief Information Officer at Santam give consent in principle to allow Zilungile Bokolo, a student at the Cape Peninsula University of Technology, to collect data in this company as part of her MICT research. The student has explained to me the nature of her research and the nature of the data to be collected.

This consent in no way commits any individual staff member to participate in the research, and it is expected that the student will get explicit consent from any participants. I reserve the right to withdraw this permission at some future time.

In addition, the company's name may or may not be used as indicated below. (Tick as appropriate.)

	Thesis	Conference paper	Journal article	Research poster
Yes				
No	x	x	x	x

Yours sincerely



Kevin Wright

11/05/2021

Date

Request to conduct research and interview participation consent letter

Dear Kevin,

Zilungile Bokolo is registered for the MICT: Master of Information Communication Technology degree at CPUT (215296273). The thesis is titled: ***Data Security in Chatbots for the Insurance Industry: A case study of a South African Insurance Company***. The researcher would like to request a permission to conduct this research at Santam insurance. The aim of this research is to develop a threat model for data security in chatbots, which can be used in practice by South African insurance organizations.

To meet the requirements of the university's Higher Degrees Committee (HDC) the student must get consent to collect data from organisations which they have identified as potential sources of data. In this case the student will use semi-structured interviews to gather data.

employees will not be forced or deceived to participate in the study. A consent letter will be signed by all participants. Participants will be free to withdraw from the study at any time that they choose. The confidentiality will be maintained and restrict illegal access to the study data. The data gathered from interview transcripts, and notes will be stored in secure storage devices. Sensitive personal data that pertain to company or individuals will be de-identifier to protect the identity of stakeholders.

The interview will be one on one and will be conducted over the period of study. Due to covid-19 restriction the interview will be conducted via teams. The targeted stakeholders are as follow, 1 Security architecture, 2 Security specialist, 2 participants from Red team, 3 participants from GTI and 1 participant from BIS.

If you agree to this, you are requested to complete the attached form and print it on your organisation's letterhead.

For further clarification on this matter please contact either the supervisor(s) identified below, or the Faculty Research Ethics Committee secretary (Ms V Naidoo) at 021 469 1012 or naidoove@cput.ac.za.

Yours sincerely,

Zilungile Bokolo
Researcher/Masters Student (CPUT)
Department of Informatics and Design
Faculty of Information Technology
Peninsula University of Technology
Email: Bokolo.zilungile348@gmail.com



Prof Justine Olawande Daramola
Research Supervisor
Department of Informatics and Design
Faculty of Information Technology
Peninsula University of Technology
Email: Daramolaj@cput.ac.za



INFORMED CONSENT FORM FOR A RESEARCH PROJECT

The consent form should be read and understood by the researcher and the participant should answer the questions asked. Before participating in the interview, the researcher and the participant should sign two copies of this form. The participant will be given one copy of the signed form.

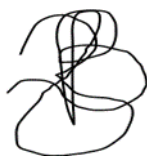
Consent Form for Participation in Interview Research

I volunteer to participate in a research project conducted by Z. Bokolo (215296273) from the Information Technology Department. I understand that the project is designed to gather information about **Chatbot in a South African Insurance Organisation**.

1. My participation in this project is voluntary. I understand that I will not be paid for my participation. I may withdraw and discontinue participation at any time without penalty. If I decline to participate or withdraw from the study, no one in my department will be told.
2. I understand that the researcher will not identify me by name in any reports using information obtained from this interview and that my confidentiality as a participant in this study will remain secure. Subsequent uses of records and data will be subject to standard data use policies which protect the anonymity of individuals and institutions.
4. I have read and understood the explanation provided to me. I have had all my questions answered me, and I voluntarily agree to participate in this study.

Participant

Date




Researcher

Date

APPENDIX D: PROFESSIONAL EDITOR'S CERTIFICATE

Appendix D aimed to show a professional editor's certificate document from a professional and accredited editor of the Cape Peninsula University of Technology that edited this thesis.



University of Pretoria
Faculty of Humanities
Department of English

This is to certify that

AA Ekata

has successfully completed the

**Programme on
Editing Principles and Practices**

1 February to 20 June 2005

Noomé
Course Leader

awulf
Operations Manager: CE at UP

92851 2008/10/06 ID: A1228740 P000293-01-2005

Handwritten Certificate Text:
EK BEWEEK DAT HIERDE DOKUMENT 'N WAARE AFSKRYF IS VAN DIE OORSPRONKELIKE DOKUMENT WAT AAN MY VERKRY IS. DIE OORSPRONKELIKE DOKUMENT DAT, VOLGENS MY WAAKSEMERS, DAAR WEE 'N WYSGING OF VERANDERING OP DIE OORSPRONKELIKE DOKUMENT ANGEWYSGING IS. I CERTIFY THAT THIS DOCUMENT IS A TRUE REPRODUCTION (COPY) OF THE ORIGINAL DOCUMENT WHICH WAS HANDED TO ME FOR AUTHENTICATION. I FURTHER CERTIFY THAT, FROM MY OBSERVATIONS, AN AMENDMENT OR A CHANGE WAS NOT MADE TO THE ORIGINAL DOCUMENT.

Signature: *MR. M. M. M. M.*
HANDTEKENING/SIGNATURE: *MR. M. M. M. M.*
RANG: *MR. M. M. M. M.*
MAGNUMMER: *MR. M. M. M. M.*
FORCE NUMBER: *MR. M. M. M. M.*
NAAM IN DRUKSKRYF: *MR. M. M. M. M.*
NAME IN PRINT: *MR. M. M. M. M.*

Stamp:
SOUTH AFRICAN POLICE SERVICE
GARISFONTEIN
2023-04-06
COMMUNITY SERVICE CENTRE
SUID-AFRIKAANSE POLISIEDIENS