Developing an intelligent online counselling framework for an employee health and wellness programme

I Kanes



orcid.org/0000-0001-6826-2136

Thesis accepted in fulfilment of the requirements for the degree Doctor of Philosophy in Social Work at the

North-West University

Promoter: Prof WJH Roestenburg

Co-promoter: Prof H Malan

Graduation: June 2025

NWU Higher Degrees Administration

SOLEMN DECLARATION AND PERMISSION TO SUBMIT

1. Solemn declaration by student

I, Ivan Kanes

declare herewith that the thesis/dissertation/mini-dissertation entitled (exact registered/approved title),

Developing an intelligent online counselling framework for an employee health and wellness programme

which I herewith submit to the North-West University is in compliance/partial compliance with the requirements set for the degree:

Doctor of Philosophy in Social Work

is my own work, has been text-edited in accordance with the requirements and has not already been submitted to any other university.

LATE SUBMISSION: If a thesis/dissertation/mini-dissertation of a student is submitted after the deadline for submission, the period available for examination is limited. No guarantee can therefore be given (should the examiner reports be positive) that the degree will be conferred at the next applicable graduation ceremony. It may also imply that the student would have to re-register for the following academic year.

Ethics number: NWU-00178-22-A1 ORCiD: 0000-0001-6826-2136

Signature of Student: Ivan Kanes Digitally signed by Ivan Kanes Date: 2024.10.27 16:17:08 +02'00' University Number: 20728670

Signed on this 27 day of October of 20 24

2. Permission to submit and solemn declaration by supervisor/promoter

The undersigned declares that the thesis/dissertation/mini-dissertation:

- ✓ Complies with the A-rules and the technical requirements provided for in the Manual for Higher Degree studies and in faculty rules;
- ✓ Has been checked by me for plagiarism (by making use of TurnItIn software for example) and a satisfactory report has been obtained, and;
- ✓ That the work was language edited before submission for examination.

Faculty specific requirements as per A-rules: 1.3.2, 4.3.2, 4.3.3, 4.4.2, 4.10.4, 5.3.2, 5.4.2, 5.10.2

- ✓ Complies with regards to faculty rules on submission or acceptance by an accredited scientific journal;
- ✓ Complies with regards to faculty rules on peer reviewed conference proceedings;
- ✓ The student is hereby granted permission to submit his/her mini-dissertation/ dissertation/thesis for examination.

Signatures of supervisor(s) and Promoter(s): (only compulsory in cases where there are co- or assistant- supervisor(s/promoters)

Willem Roestenburg Digitally signed by Willem Roestenburg Date: 2024.10.29 11:12:44 +02'00' Promoter

Prof Hanelie Malan Digitally signed by Prof Hanelie Malan Date: 2024.10.29 13:28:55 +02'00' Co-Supervisor/ Co-Promoter

Assistant -Supervisor Assistant-Promoter

Original details: (10512187) M:\HDA Toolbox\Forms\Pdf fillable forms\13.Solemn Decl and Permission to submit exam copy.pdf 3 December. File reference: 7.1.11.3.2/3

DEDICATION

This thesis is dedicated to my sister, Charlene.

Because you made it, I was able to make it.

PREFACE

One would never think that a simple interview question would lead to the completion of a thesis. During an interview for a position at my current company, I was asked a pertinent question: “What are your thoughts on the Fourth Industrial Revolution (known as the 4IR), and is South Africa equipped to handle it?” I had to ask what 4IR even means - needless to say, I completely botched the question. This question did, however, pave the way for my interest in everything 4IR-related. My readings prompted me to propose the current study.

For many years, popular literature and scientists have been occupied with the future and machines that could do what humans could do or improve human capacity. The first real breakthrough was Marian Rejewski, Jerzy Różycki, and Henryk Zygalski’s Bomba, which managed to decipher the German’s Enigma machine during World War II (WWII). While Alan Turning did not directly work on Bomba, his contributions were also significant in code-breaking efforts. To this day, thinking machines are subjected to the Turning Test, which brings us to our topic.

We find ourselves in an exciting period of history where artificial intelligence (AI) is rapidly evolving and becoming part of our daily lives. Access to and the use of AI is no longer confined to a select few scientists and corporations – it has become available to every person who has access to the Internet. When referring to AI, the Chat Generative Pre-Trained Transformer (ChatGPT) and the recently-added Meta AI immediately come to mind. AI is not limited to these handy tools that we have at our disposal but reaches into sectors ranging from agriculture to healthcare. The use of AI in various sectors prompted the question: How are we using AI in mental healthcare and more specifically social work?

This study investigated the live chat counselling service offered by one of the most prominent EAP providers in Africa. Live chat counselling services are conducted by human mental healthcare practitioners, who provide psychological therapy services to their clients via a text-based platform.

This study not only aimed to investigate the reasons for which clients would engage in live chat counselling sessions but also aimed to understand how they conveyed their concerns or messages. To investigate the aforementioned, a unique approach to social work research methodology was developed. This methodology deviates from the more popular forms of social work research in that a multi-phased mixed-method approach was followed. The approach consisted of a ‘qualitative first’, ‘quantitative second’ approach, incorporating thematic-, word frequency- and sentiment analyses.

Certain insights were gained during the empirical phase of the study, whereafter findings indicated that an AI generative chatbot based on big data would be a viable option for social work to remain relevant in a rapidly evolving technological landscape.

Venturing into the depths of mental healthcare, this study delves into the cutting-edge technologies currently revolutionising the field, where AI becomes the beacon guiding our exploration of innovative solutions to pressing mental health challenges. The investigation uncovered numerous psychological entities leveraging AI to deliver transformative mental healthcare services. These services are, however, focused on certain psychological paradigms. Extensive research confirmed that services such as this do not exist in social work.

This study, based on empirical evidence and sound theoretical frameworks, proposes an intelligent online counselling framework for an employee health and wellness programme (EHWP).

ACKNOWLEDGEMENTS

I would like to extend my gratitude to Lyra Southern Africa for providing access to their live chat data and their extensive affiliate network.

ABSTRACT

The Fourth Industrial Revolution (4IR), with its transformative technological advancements, particularly in artificial intelligence (AI) is shaping various sectors, including mental healthcare. This thesis investigated the development of an intelligent online counselling framework for a social work employee health and wellness programmes (EHWP), specifically within the field of social work. The research investigated the role of AI, especially generative chatbot use and how it can supplement human counsellors through text-based mental health support.

By drawing upon a multi-phased mixed-method approach, the study integrated qualitative and quantitative analyses of live chat counselling sessions. The study evaluated existing AI applications in mental health, identified trends in live chat counselling transcripts and explored a conceptual AI framework informed by empirical findings.

The study commenced with a literature study, which examined the current use of AI in mental healthcare and its potential application in social work. The literature study formed the foundation for the empirical phase of the research. The qualitative phase of the research employed thematic analysis to explore prominent themes and trends in live chat sessions, while the quantitative phase utilised lexicon-based analysis techniques to examine the frequency and intensity of these themes. By converging these findings, the research developed a nuanced understanding of the challenges and opportunities presented by an intelligent online counselling framework for a social work EHWP.

The proposed framework incorporated advanced AI technologies such as natural language processing (NLP) and machine learning (ML) to create a generative online counselling framework capable of understanding and responding to a range of mental health needs. The framework has been proposed as a supplementary function for human counsellors, capable of providing immediate, consistent and empathetic support to users.

A Delphi panel consisting of social workers, registered counsellors and psychologists was convened to critically review and evaluate the proposed product. The Delphi panellists' feedback and recommendations were analysed and subsequently incorporated into the proposed framework, thereby improving the framework's ability to meet the psychosocial needs of potential users.

The research aimed to contribute towards the modernisation of social work through the integration of AI technologies and ultimately proposed a viable innovative approach to employee health and wellness through the combination of technological efficiency with human empathy. This

framework aims to improve service delivery by fostering multidisciplinary collaboration and future research in AI applications in social work.

The findings suggested that AI could play a significant role in enhancing the accessibility and effectiveness of mental health services, especially in high-demand contexts, such as EHWPs. By addressing practical and psychosocial needs, the intelligent online counselling framework represents a significant step towards the integration of AI into social work, with the prospect of transforming the delivery and experience of mental health services.

Key terms: Artificial Intelligence (AI); Chatbot; Health information technology; Mental health; Social work

MANUSCRIPTS SUBMITTED FOR PUBLICATION

Kanes, I., Roestenburg, W.J.H. & Malan, H. 2024. Developing an intelligent online counselling framework for an employee health and wellness programme: a literature review. *Social Work/Maatskaplike Werk.* Manuscript submitted for review and publication.

Kanes, I., Roestenburg, W.J.H. & Malan, H. 2024. Transforming Mental Health Support: AI Chatbots and Insights from Live Chat. *International Social Work.* Manuscript submitted for review and publication.

Kanes, I., Roestenburg, W.J.H. & Malan, H. 2024. An intelligent online counselling framework for an employee health and wellness programme. *International Social Work*. Manuscript submitted for review and publication.

PRESENTATIONS

Kanes, I., Roestenburg, W.J.H. & Malan, H. 2023. Developing an intelligent online counselling framework for an employee health and wellness programme. ASASWEI Conference, September 2023, Durban, South Africa.

LIST OF ACRONYMS

4IR	Fourth Industrial Revolution
AI	Artificial Intelligence
AI4SG	Artificial Intelligence for Social Good
AIML	Artificial Intelligence Markup Language
ANN	Artificial Neural Networks
App	Application
AUDIT	Alcohol Use Disorders Identification Test
BERT	Bidirectional Encoder Representations from Transformers
BSPS	Brief Social Phobia Scale
BTQ	Brief Trauma Questionnaire
CAQDAS	Computer-Assisted Qualitative Data Analysis Software
CBT	Cognitive Behavioural Therapy
ChatGPT	Chat Generative Pre-Trained Transformer
COMPRES	Community Psychosocial Research
DL	Deep Learning
DM	Dialogue Management
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
EAP	Employee Assistance Program
EHWP	Employee Health and Wellness Programme

FNE	Fear of Negative Evaluation Scale
GAD-7	Generalised Anxiety Disorder 7-item
GCS	Generalised Contentment Scale
GPP	Genuine People Personalities
HANDS	Harvard Department of Psychiatry/NDSD scale
HCI	Human-Computer Interaction
HDRS	Hamilton Depression Rating Scale
HITs	Health Information Technologies
HPCSA	Health Professions Council of South Africa
HREC	Health Research Ethics Committee
HSSF	Heimler Scale of Social Functioning
IBM	International Business Machines
ICAS	Independent Counselling and Advisory Service
ICT	Information and Communication Technology
IoT	Internet of Things
IT	Information Technology
LSTM	Long-Short-Term Memory
MAST	Michigan Alcoholism Screening Test
MBCT	Mindfulness-Based Cognitive Therapy
ML	Machine learning
NASW	National Association of Social Workers
NER	Named Entity Recognition

NLG	Natural Language Generation
NLP	Natural Language Processing
NLTK	Natural Language Toolkit
NLU	Natural Language Understanding
NWU	North-West University
PHQ-9	Patient Health Questionnaire-9
POS	Part-of-Speech
PST	Problem-Solving Therapy
PSWQ	Pennsylvania State Worry Questionnaire
PSWQ	Penn State Worry Questionnaire
PTSD	Post-Traumatic Stress Disorder
SACSSP	South African Council for Social Service Professions
SADS	Social Avoidance and Distress Scale
SAHPRA	South African Health Products Regulatory Authority
SDLC	Software Development Life Cycle
SFBT	Solution-focused brief therapy
SPSS	Statistical Package for the Social Sciences
SSWE	Sentiment-Specific Word Embedding
USD	United States Dollar
XML	Extensible Markup Language
WWII	World War II

DEFINITIONS

Artificial intelligence

No standard definition of artificial intelligence (AI) exists. In its most simplistic description, AI can be regarded as a machine displaying intelligent behaviour (intelligence being a point of contention in itself) where complex tasks can be performed. Even though there is no single definition of AI due to its many characteristics, definitions reveal several similarities. These common characteristics are expressed in a way that suggests they might be regarded as the primary characteristics of AI, which include i) the perception of the environment, taking into account the complexity of the real world, ii) processing information (gathering and analysing inputs), iii) making decisions (including thinking and learning) iv) acting and carrying out activities with some degree of autonomy (including adaptation and responding to environmental changes), and v) the achievement of certain objectives being the primary motivation for artificially intelligent systems (Ertel, 2017; Kaplan & Haenlein, 2019; Samoili *et al.*, 2020).

The following examples illustrate this definition (George *et al.*, 2021; Giachos *et al.,* 2023):

- Virtual assistants such as Siri, Alexa and Cortana draw upon the natural language process to interpret and respond to spoken commands;
- Self-driving cars draw upon machine learning (ML) algorithms to recognise and respond to traffic patterns and road conditions;
- Fraud detection systems use data analysis to identify patterns of fraudulent behaviour; and
- Personalised recommendation systems, such as those used by Netflix and Amazon, analyse user data and recommend content based on individual preferences.

Chatbot

With the development and analysis of intelligent hardware and software, also known as intelligent agents, AI is gradually integrating into our daily lives. Intelligent agents can perform a range of tasks, from manual labour to complex procedures. One of the simplest and most common forms of intelligent human-computer interaction (HCI), a chatbot, is a basic example of an AI system. Chatbots are online computer programs that serve as virtual assistants in social networks or web applications by conversing with users in natural language, which includes voice or text language. Chatbots can vary in complexity, from simple rule-based systems to advanced ML models that can learn and adapt over time (Adamopoulou & Moussiades, 2020; Perez-Soler *et al.*, 2021).

Some popular chatbots include:

- Endurance: A dementia patient's companion (Pathan *et al.*, 2021);

- Casper: Assisting insomniacs through the night (Shaikh *et al.*, 2019);
- Chat Generative Pre-Trained Transformer (ChatGPT): An advanced language model that uses AI to generate human-like responses in conversations (OpenAI, 2023); and
- MedWhat: Accelerating medical diagnosis (Bulla *et al.,* 2020).

Framework

In the context of this study, a framework is an organised method containing specifications for system development. By employing a variety of methods, it forms the basis for the development and integration of various components in the overall design of an AI generative chatbot (Kumar & Ali, 2020).

Model

In the context of this study, a model typically refers to a representation, simulation, process, concept or system that aims to predict or describe behaviour. Models are deployed to understand and predict patterns, behaviours and responses of user interactions with the AI systems. Models can include algorithms, decision-making processes and ML techniques (Sarker *et al.*, 2021).

Tool(s)

Tools refer to numerous digital and technological applications and systems that are utilised to enhance social work practice and the delivery of social work services. These tools facilitate the effective development and deployment of user-friendly digital interventions, whereby the overall efficiency of social work services is enhanced within the context of the Fourth Industrial Revolution (4IR) (Pink *et al.*, 2022).

REFERENCES

Adamopoulou, E. & Moussiades, L. 2020. *An overview of chatbot technology*. Paper delivered at IFIP International Conference on Artificial Intelligence Applications and Innovations, Cham. https://link.springer.com/chapter/10.1007/978-3-030-49186-4_31 Date of access: 18 Dec. 2022.

Bulla, C., Parushetti, C., Teli, A., Aski, S. & Koppad, S. 2020. A review of AI based medical assistant chatbot. *Res Appl Web Dev Des*, 3:1-14.

Ertel, W. 2017. *Introduction to Artificial Intelligence*. 2nd ed. Gewerbestrasse: Springer.

George, N., Muiz, K., Whig, P. & Velu, A. 2021. Framework of Perceptive Artificial Intelligence using Natural Language Processing (PAIN). *Artificial & Computational Intelligence,* 2(2):1-8.

Giachos, I., Papakitsos, E.C., Savvidis, P. & Laskaris, N. 2023. Inquiring Natural Language Processing Capabilities on Robotic Systems through Virtual Assistants: A Systemic Approach. *Journal of Computer Science Research*, 5(2):28-36.

Kaplan, A. & Haenlein, M. 2019. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1):15-25.

Kumar, R. & Ali, M.M. 2020. A review on chatbot design and implementation techniques. *Int. J. Eng. Technol*, 7(11):2791-2800.

OpenAI. 2023. Join us in shaping the future of technology. https://openai.com/ Date of access: 28 Jan. 2023.

Pathan, H., Mansuri, I. & Pinjari, K. 2021. AI Meets Mental Health. *International Research Journal of Engineering and Technology,* 8(5):3715-3718.

Perez-Soler, S., Juarez-Puerta, S., Guerra, E. & de Lara, J. 2021. Choosing a Chatbot Development Tool. *IEEE Software*, 38(4):94-103.

Pink, S., Ferguson, H. & Kelly, L. 2022. Digital social work: Conceptualising a hybrid anticipatory practice. *Qualitative Social Work*, 21(2):413-430.

Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F. & Delipetrev, B. 2020. *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence*. Sevilla: Publications Office of the European Union.

Sarker, I.H., Khan, A.I., Abushark, Y.B. & Alsolami, F. 2021. Mobile expert system: exploring context-aware machine learning rules for personalized decision-making in mobile applications. *Symmetry*, 13(10):1975-1985.

Shaikh, A., More, D., Puttoo, R., Shrivastav, S. & Shinde, S. 2019. A survey paper on chatbots. *International Research Journal of Engineering and Technology (IRJET)*, 6(4):2395-0072.

TABLE OF CONTENTS

SOLEMN DECLARATION AND PERMISSION TO SUBMIT I

DEDICATION II

PREFACE III

ACKNOWLEDGEMENTS V

ABSTRACT VI

MANUSCRIPTS SUBMITTED FOR PUBLICATION VIII

PRESENTATIONS IX

LIST OF ACRONYMS X

DEFINITIONS XIII

LIST OF FIGURES XXII

LIST OF TABLES XXIV

LIST OF EXCERPTS XXV

NOTE TO EXAMINERS XXVI

SECTION A 1

CHAPTER 1: INTRODUCTION TO THE RESEARCH 1

1.1. TITLE 1
1.2. KEYWORDS 1
1.3. EXECUTIVE SUMMARY 1
1.4. CONTEXTUALISATION AND PROBLEM STATEMENT 1
1.4.1. Contextualisation **2**
1.4.2. Problem statement **7**
1.5. RESEARCH QUESTION 7
1.5.1. Research sub-questions **7**
1.5.2. The expected contribution of the study **8**
1.5.3. Research statement **8**
1.5.4. Proposed research product **8**
1.6. AIM AND OBJECTIVES 8
1.6.1. Aim **8**
1.6.2. Objectives **8**
1.7. OVERALL RESEARCH APPROACH 9
1.7.1. Phase 1: Literature Study **10**
1.7.1.1. Description 10

1.7.1.2. Procedure ... *11*
1.7.1.3. Data analysis methods in literature studies ... *12*
1.7.2. Phase 2: Pilot Study ... **13**
1.7.2.1. Description ... *13*
1.7.2.2. Procedure ... *13*
1.7.2.3. Sampling method ... *13*
1.7.2.4. Data analysis ... *14*
1.7.3. Phase 3: Data Collection and Analysis ... **15**
1.7.3.1. Qualitative live chat transcript collection and analysis ... *15*
1.7.3.2. Data gathering process ... *18*
1.7.3.3. Validity, reliability, and trustworthiness ... *20*
1.7.3.4. Quantitative textual data analysis ... *21*
1.7.3.5. Risks and benefits analysis ... *25*
1.7.3.6. Confidentiality, privacy, and anonymity ... *25*
1.7.4. Phase 4: Design and Development of Proposed Product ... **26**
1.7.4.1. Description ... *26*
1.7.4.2. Procedure ... *30*
1.7.5. Phase 5: Delphi Study ... **31**
1.7.5.1. Description ... *32*
1.7.5.2 Procedure ... *32*
1.7.5.3. Population ... *33*
1.7.5.4. Sampling method ... *34*
1.7.5.5. Sample size and motivation ... *35*
1.7.5.6. Data gathering process ... *35*
1.7.5.7. Validity, reliability, and trustworthiness ... *35*
1.7.5.8. Data analysis methods ... *36*
1.7.5.9. Risk level descriptors ... *36*
1.7.5.10. Risks and benefits ... *36*
1.7.5.11. Further ethical considerations ... *37*
1.8. RESEARCH BUDGET ... 37
1.9. TIMEFRAME ... 38
1.10. CHOICE AND STRUCTURE OF REPORT ... 38
1.11. EXPERTISE OF RESEARCHERS ... 39
1.12. RISK LEVEL OF THE STUDY ... 39
1.13. CONFLICT OF INTEREST ... 39
1.14. DISSEMINATION OF RESULTS ... 39
REFERENCES ... 40
CHAPTER 2: LITERATURE STUDY ... 49
2.1. INTRODUCTION ... 49
2.2. LITERATURE STUDY METHODOLOGY ... 51

2.3. RATIONALE FOR THE STUDY 51
2.4. ARTIFICIAL INTELLIGENCE (AI) CHATBOTS 52
2.4.1. Benefits of Artificial Intelligence (AI) chatbots and virtual therapeutic services **54**
2.4.2. Limitations of Artificial Intelligence (AI) chatbots and virtual therapeutic services **55**
2.4.3. Holistic features of Artificial Intelligence (AI) chatbots and virtual mental health therapeutic services **57**
2.5. METHODOLOGIES DEPLOYED IN THE DEVELOPMENT OF AI CHATBOTS 60
2.5.1. Characteristics for consideration in the design of an Artificial Intelligence (AI) chatbot framework **60**
2.5.2. Text analysis **61**
2.5.3. Data pre-processing **61**
2.5.4. Input embedding **62**
2.5.5. Natural Language Processing (NLP) **63**
2.5.6. Naïve Bayes Method **64**
2.5.7. Machine learning (ML) **66**
2.5.8. Sentiment analysis **67**
2.5.9. Artificial Neural Networks (ANN) **68**
2.5.10. Integration of research methodology in the facilitation of chatbot design **69**
2.6. THE USE OF BIG DATA IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE (AI) CHATBOTS 70
2.7. THE USE OF LINGUISTICS IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE (AI) CHATBOTS 72
2.8. THE USE OF AI CHATBOTS WITHIN A SOCIAL WORK AND AN EHWP CONTEXT 73
2.9. CLINICAL THEORETICAL FRAMEWORK 76
2.10. ETHICAL CONSIDERATIONS IN THE DEPLOYMENT OF AI IN A SOCIAL WORK EHWP 77
2.11. RECOMMENDATIONS 78
2.12. CONCLUSION 80
REFERENCES 82

CHAPTER 3: EMPIRICAL FINDINGS 94

3.1. INTRODUCTION 94
3.2. RATIONALE FOR STUDY 94
3.3. CURRENT USE OF AI IN SOCIAL WORK AND EHWP 95
3.4. METHODOLOGY AND RESEARCH DESIGN 96
3.5. EMPIRICAL DATA: LIVE CHAT TRANSCRIPT ANALYSIS 97
3.5.1. Theme 1: Flow/process flow of live chat sessions **97**
3.5.2. Theme 2: Problems/challenges **99**
3.5.2.1. Sub-theme 2.1. Personal development: General 103
3.5.2.2. Sub-theme 2.2. Dysfunction 104
3.5.2.3. Sub-theme 2.3. Psychological challenges: Uncategorised 106
3.5.2.4. Sub-theme 2.4. Emotional wellbeing 108

3.5.2.5. Sub-theme 2.5. Career and work-related: General *109*
3.5.2.6. Sub-theme 2.6. Interpersonal relationships *111*
3.5.2.7. Sub-theme 2.7. Family dynamics *113*
3.5.2.8. Sub-theme 2.8. Request for help/support *114*
3.5.2.9. Sub-theme 2.9. Suicidal ideation *116*
3.5.2.10. Sub-theme 2.10. Diverse challenges/issues/problems *117*
3.5.3. Theme 3: Counsellor approaches **118**
3.5.3.1. Sub-theme 3.1. Advice and support *119*
3.5.3.2. Sub-theme 3.2. Psychological approach *121*
3.5.3.3. Sub-theme 3.3. Therapeutic approach *123*
3.5.4. Theme 4: Linguistic trends **125**
3.5.5. Theme 5: Technical issues **127**
3.6. SENTIMENT ANALYSIS: LIVE CHAT TRANSCRIPT ANALYSIS 129
3.6.1. Negative sentiments **130**
3.6.2. Neutral sentiments **130**
3.6.3. Positive sentiments **131**
3.7. TEXTUAL TRENDS 131
3.8. DISCUSSION OF RESULTS AND DATA CONVERGENCE 133
3.9. RECOMMENDATIONS 135
3.10. CONCLUSION 136
REFERENCES 137

CHAPTER 4: PROPOSED PRODUCT: AN INTELLIGENT ONLINE COUNSELLING FRAMEWORK FOR AN EMPLOYEE HEALTH AND WELLNESS PROGRAMME 142

4.1. INTRODUCTION 142
4.2. DELPHI PANEL METHODOLOGY FOR REFINING THE AI COUNSELLING FRAMEWORK 143
4.3. THE REFINED INTELLIGENT ONLINE COUNSELLING FRAMEWORK 145
4.3.1. Planning phase **146**
4.3.1.1. Goal 1: Accessible counselling *146*
4.3.1.2. Goal 2: Address user-specific concerns/challenges/problems *147*
4.3.1.3. Goal 3: Evidence based assessments *147*
4.3.1.4. Goal 4: Draw upon a sound psychological theoretical framework *147*
4.3.1.5. Goal 5: Draw upon sound counselling approaches *147*
4.3.1.6. Goal 6: Understand and mimic linguistic styles *147*
4.3.1.7. Goal 7: Recognise and interpret triggers words *148*
4.3.2. ANALYSIS PHASE **148**
4.3.2.1. Identification of problems/needs *148*
4.3.2.2. Functional requirements *149*
4.3.3. AN INTELLIGENT ONLINE COUNSELLING FRAMEWORK: CLINICAL TECHNICAL REQUIREMENTS **149**
4.3.3.1. Natural Language Processing (NLP) *149*

4.3.3.2. Emotion recognition and affective computing *152*
4.3.3.3. Contextual awareness *154*
4.3.3.4. Knowledge base integration *157*
4.3.3.5. Dynamic risk assessment and crisis response *160*
4.3.3.6. Adaptability and learning *163*
4.3.3.7. Therapeutic approach, goal setting and progress tracking *165*
4.3.3.8. Referral mechanism *169*
4.3.3.9. Cultural sensitivity *170*
4.3.3.10. Personalisation *172*
4.3.3.11. Empathy and tone *173*
4.3.3.12. User engagement strategies *175*
4.3.3.13. Community resources or support groups *177*
4.3.3.14. Evidence-based assessments *178*
4.3.3.15. Bayes' theorem *181*
4.3.3.16. User feedback and improvement mechanism *183*
4.3.3.17. Ethical integrity in the Artificial Intelligence (AI) counselling chatbot architecture *184*
4.3.3.18. Integration with mobile application *185*
4.3.3.19. Integration with administrative system *186*
4.4. CONCLUSION 187
4.5. RECOMMENDATIONS 189
4.5.1. Implementation recommendations **189**
4.5.2. Operational recommendations **189**
4.5.3. Future research opportunities **189**
REFERENCES 191

CHAPTER 5: FINDINGS SUMMARY, LIMITATIONS OF THE STUDY, CONCLUSION AND RECOMMENDATIONS 197

5.1. SUMMARY 197
5.1.1. Chapter 1: Introduction to the research **197**
5.1.2. Chapter 2: Literature Study **198**
5.1.3. Chapter 3: Empirical Findings **200**
5.1.4. Chapter 4: Proposed Product: An Intelligent Online Counselling Framework for an Employee Health and Wellness Programme (EHWP) **201**
5.2. LIMITATIONS OF THE STUDY 202
5.2.1. Subjectivity in thematic analysis **202**
5.2.2. Adapting the research focus **202**
5.2.3. Limited literature on AI in social work **203**
5.3. CONCLUSION 203
5.4. RECOMMENDATIONS 204
5.4.1. Practical applications **204**
5.4.2. Policy recommendations **205**

5.4.3. Future research opportunities **205**

SECTION B 207

COMPLETE LIST OF REFERENCES 207

ANNEXURES 235

ANNEXURE 1: SCIENTIFIC COMMITTEE APPROVAL 235
ANNEXURE 2: LETTER OF AUTHORISATION TO CONDUCT RESEARCH AT ICAS SOUTHERN AFRICA 237
ANNEXURE 3: LETTER OF AUTHORISATION TO CONDUCT RESEARCH AT LYRA SA – ACCESS TO AFFILIATE NETWORK 238
ANNEXURE 5: INFORMED CONSENT DOCUMENTATION FOR: HEALTH INFORMATICS EXPERT AND MENTAL HEALTH CARE PROFESSIONAL DELPHI PANELISTS 244
ANNEXURE 6: EDITORIAL LETTER 249

LIST OF FIGURES

Figure 1. Example of a live chat interaction....6
Figure 2. Overall research approach....10
Figure 3. Example of consent (counsellor's view)....18
Figure 4. Example of consent (service user view)....19
Figure 5. Example of text analysis....22
Figure 6. The architecture of a typical chatbot....28
Figure 7. Proposed product concept design/process flow....30
Figure 8. Waterfall model of the SDLC....31
Figure 9. Gantt Chart of timeframe....38
Figure 10. AI chatbots versus traditional modes of communication....58
Figure 11. Commonly deployed strategies in chatbot framework development....61
Figure 12. Example of a simple ANN....69
Figure 13. Integration of IT into social work and EHWPs....75
Figure 14. Process flow of live chat sessions....99
Figure 15. Count of code groups....100
Figure 16. Problem network....102
Figure 17. Sub-theme 2.1. Personal development: General network....104
Figure 18. Sub-theme 2.2. Dysfunction network....106
Figure 19. Sub-theme 2.3. Psychological challenges: Uncategorised network....107
Figure 20. Sub-theme 2.4. Emotional wellbeing network....109
Figure 21. Sub-theme 2.5. Career and work-related: General network....111
Figure 22. Sub-theme 2.6. Interpersonal relationships network....113
Figure 23. Sub-theme 2.7. Family dynamics network....114
Figure 24. Sub-theme 2.8. Request for help/support network....115
Figure 25. Sub-theme 2.9. Suicidal ideation network....117
Figure 26. Counsellor approaches....119
Figure 27. Advice and support....120
Figure 28. Psychological approach....123
Figure 29. Therapeutic approach....124
Figure 30. Linguistic trends....126
Figure 31. Sentiment analysis....129
Figure 32. Live chat session flow and elements....134
Figure 33. Componential model....153
Figure 34. Contextual awareness....155
Figure 35. Dynamic risk assessment and crisis response....162
Figure 36. Three-phase model....175

Figure 37. Evidence-based assessments and follow-up 179
Figure 38. Integration with mobile application 186
Figure 39. Integration with administrative system 187
Figure 40. Conceptual summary of an intelligent online counselling framework 189

LIST OF TABLES

Table 1. Live chat transcript inclusion criteria 17
Table 2. Summary of data analysis procedure 23
Table 3. Research budget 37
Table 4. Features of mental health AI chatbots (adapted from Abd-alrazaq *et al.*, 2019) 59
Table 5. Problem/challenge groupings 100
Table 6. Occurrences of problems/challenges 100
Table 7. Counsellor approaches 118
Table 8. Therapeutic approach 125
Table 9. Linguistic trends 127
Table 10. Technical 128
Table 11. Sentiment analysis 129
Table 12. Textual trends 132
Table 13. Delphi panellist expertise, affiliations, and key contributions 144
Table 14. Linguistic styles 151
Table 15. Empirical findings of sentiment analysis 152
Table 16. Emotive expressions occurring >0.01% 153
Table 17. Reasons for user engagement 157
Table 18. Risk codes 160
Table 19. Risk trigger words 162
Table 20. Problem clusters 165
Table 21. Therapeutic approaches 167
Table 22. Social and environmental factors 171
Table 23. Evidence-based assessments 179
Table 24. Ethical considerations 184
Table 25. AI ethical principles 184
Table 26. Integration with administrative system 186

LIST OF EXCERPTS

Excerpt 1. Personal development: General 103
Excerpt 2. Dysfunction 105
Excerpt 3. Psychological challenges: Uncategorised 107
Excerpt 4. Emotional wellbeing 108
Excerpt 5. Career and work-related: General 110
Excerpt 6. Interpersonal relationships 112
Excerpt 7. Family dynamics 114
Excerpt 8. Request for help/support 115
Excerpt 9. Suicidal ideation 116
Excerpt 10. Diverse challenges/issues/problems 118
Excerpt 11. Advice and support 120
Excerpt 12. Psychological approach 122
Excerpt 13. Therapeutic approach 124
Excerpt 14. Linguistic trends 126
Excerpt 15. Negative sentiments 130
Excerpt 16. Neutral sentiments 131
Excerpt 17. Positive sentiments 131
Excerpt 18. Example of user linguistic style 152
Excerpt 19. User emotive expressions 154
Excerpt 20. Contextual awareness 156
Excerpt 21. Need for extensive knowledge base 159
Excerpt 22. Risk engagement 163
Excerpt 23. Adaptability and learning 164
Excerpt 24. Use of SFBT techniques 169
Excerpt 25. User request for referral 170
Excerpt 26. Social and environmental factors 172
Excerpt 27. Personalisation 173
Excerpt 28. Empathy and tone 174
Excerpt 29. Three-phase model 177
Excerpt 30. Community resources or support groups 178
Excerpt 31. Evidence-based assessment 180
Excerpt 32. Illustration of the use of Bayes' theorem 183
Excerpt 33. User feedback 183

NOTE TO EXAMINERS

Examiners, please note the following:

1. This submission has been submitted in the form of a thesis. Section A includes:
    a. Chapter 1: Research methodology;
    b. Chapter 2: Literature study;
    c. Chapter 3: Empirical findings;
    d. Chapter 4: Proposed product; and
    e. Chapter 5: Findings summary, limitations, conclusion and recommendations.
2. Page numbering: For the sake of ease in reading, page numbering in this thesis is consecutive throughout.
3. Section B: A comprehensive reference list is provided at the end of the thesis encompassing all sources cited throughout the chapters and articles. The reference list follows the citation style prescribed by North-West University's (NWU) Harvard style.
4. Ethical considerations: This research adheres to ethical standards and guidelines set forth by the NWU Health Research Ethics Committee (HREC). Any ethical concerns or considerations pertaining to the research are addressed in Chapter 1: Research Methodology.
5. Access to data: Any data or materials relevant to the research are available upon request, per HREC policies on data sharing and accessibility.
6. Confidentiality and data protection: Any sensitive or confidential information obtained during the research process is handled following applicable laws, regulations, and ethical guidelines. Measures have been taken to ensure the confidentiality and privacy of participants.
7. Language and writing style: The thesis is written in English (South Africa) with attention to clarity, coherence, and academic rigour. The writing style is scholarly and formal, adhering to disciplinary conventions and standards. The excerpts are written in English (American), as that is how the transcription was performed.
8. Declarations: Declarations regarding originality, authorship, and any conflicts of interest are included in the thesis, as required by NWU's guidelines.
9. The data and ethical requirement documentation reference the Independent Counselling and Advisory Service (ICAS). It should be noted that ICAS has merged with Lyra and is now formally recognised as Lyra SA. Additionally, any references to "the host company" pertain to Lyra SA.

SECTION A

CHAPTER 1: INTRODUCTION TO THE RESEARCH

1.1. TITLE

Developing an intelligent online counselling framework for an employee health and wellness programme

1.2. KEYWORDS

Artificial Intelligence, Chatbots, Health Information Technology, Mental Health, Social Work.

1.3. EXECUTIVE SUMMARY

Since 2016, firms such as Microsoft, Facebook, and Google have recognised chatbots and conversational user interfaces as an evolving means of interacting with data and services. Research and practitioner interest in these validations has increased, resulting in developments that previously appeared unthinkable (Folstad *et al.*, 2018).

Eliza, a human-computer interface, was the first artificial intelligence (AI) agent that was developed and used in psychology. Eliza was a Carl Rogers-esque therapist who formulated responses mimicking a conversation (Ausman, 2019). Eliza opened the path for numerous other AI therapy agents, for instance, Horizons (social therapy), Woebot (cognitive behavioural therapy [CBT]), Paro (loneliness crisis aid), Second Life (speech and social therapy), Moodkit (CBT), Moodnotes (mood tracking), Pacifica (CBT) and Wysa (CBT), amongst others (De Jesus, 2019).

While several well-researched and well-developed mental health applications and agents are available in the field of psychology, a void for these agents remains in the science of social work. Tambe and Rice (2018) propose a radical shift in the social work perspective whereby they ‘dare’ social workers to imagine how AI can benefit their client populations.

This study proposes exploring and designing a human-computer interface agent or chatbot framework, in the science field of social work. This is specifically aimed at the psychosocial domain of an employee health and wellness programme (EHWP).

1.4. CONTEXTUALISATION AND PROBLEM STATEMENT

This section discusses the context and background within a social work and EHWP space. Advancements in the Fourth Industrial Revolution (4IR) continue to shape societal frameworks as technology offers new pathways and challenges for social work practice. The sub-sections

provide a contextualisation and background to the problem statement that motivates the study's aim, namely to conceptualise and AI-based generative chatbot as a supplementary tool in social work and EHWP settings.

1.4.1. Contextualisation

The way people live, work and interact with one another is currently undergoing a significant change due to technology. This shift is unlike anything humans have ever encountered in its size, scope, and complexity (Deloitte Development, 2018; Schwab, 2016). The production and retention of knowledge are shifting from human beings to online entities that can apply cognitive thinking, imagination, and creativity. As the 4IR arose, humanity underwent significant changes in the social spheres of society based on the adoption of social life through data extraction (Benyera, 2021). As a result, the 4IR differs from previous industrial revolutions in terms of velocity (evolving at a faster rate), breadth and depth (a digital revolution combining multiple technologies, resulting in paradigm shifts), and systems impact (the evolution and transformation of entire systems).

According to Schwab (2016), multi-stakeholder collaboration across disciplinary, social, political, national, and industry borders is essential. Such multifaceted cooperation will shape the 4IR to ensure it is uplifting and human-centred rather than dividing and dehumanising.

One of the driving forces behind the 4IR is the Internet of Things (IoT), which bridges physical and digital applications. In its most basic form, the IoT can be described as a relationship made possible by connected technologies and diverse platforms between commodities (for example, products and services) and people. Customers, individuals or companies are at the centre of the IoT, which aims to improve how these populations provide services. Technological innovations are transforming how organisations perceive and manage their assets with advanced digital capabilities to increase their value. In addition, new operating models require companies and organisations to rethink their operating models, to operate faster and with greater agility - though the implementation of AI is already vastly noticeable in self-driving cars, drones, and virtual assistance, amongst others (Schwab, 2016; Schwab, 2015).

Until recently, computers competed with humans mainly in physical skills and jobs; however, machines are now beginning to compete with humans in cognitive skills (Harari, 2019). Robotics is increasingly being used across various sectors, performing various functions ranging from precision agriculture to healthcare. Collaboration between people and machines will soon become commonplace thanks to rapid advancements in robotics.

As per Harari (2016), machine learning (ML) technology is already capable of performing tasks that once seemed to reflect the pinnacle of human creativity. Researchers can collect and analyse a wealth of data using ML and other new technologies, giving us access to previously

inconceivable numbers and types of information (Xu *et al.*, 2018). The difficulty lies in progressing to the next stage, which involves altering the way we train, think, and engage while employing data to generate value from the knowledge discovered by cutting-edge technology.

Health information technologies (HITs) have the potential to modernise the healthcare industry (Levinson, 2018). A study by Roestenburg (2021) mentions that information technology (IT) systems enhance and facilitate operations, which could enhance client satisfaction while consuming the least resources. Neuman *et al.* (2013) assert that the only way to control healthcare costs without compromising the quality of care in the long term is to adopt the use of HITs. Mental health chatbots and therapy applications such as HITs are still in their very early stages and need a great deal of study and streamlining to ensure that customer bases are treated appropriately, however numerous counselling bots, mainly in the discipline of psychology, exist. Platforms such as Woebot, Moodkit, Moodnotes, Pacifica, Wysa, and Joy Application (App) have been well-researched and developed by renowned companies and academic institutions (De Jesus, 2019).

Social work consists of various services offered to individuals, groups, and communities to strengthen social functioning within various contexts. Social work can be seen as the most dominant profession that focuses on social functioning (Van Ewijk, 2010). Social work as a broad field consists of various types of users, professionals, and services, which aim to support service users in participating in institutions such as families, labour, leisure, health, education, and communities.

According to Van Ewijk (2018), the social work profession is value-based in its origin. It draws on the values of compassion, justice, and care. Therefore, social work can be positioned as a clearly defined profession, an umbrella for various social professions, a discipline with various branches of practice, a field of practice, and a body of knowledge. At the basic level, the most important social work principles are self-determination, participation, empowerment, and a generic holistic approach.

In the context of this study and a more extensive EHWP, social work may be a field of practice where social workers address the human and social requirements of the workplace by creating and putting into action the proper interventions to promote healthier people and settings (Googins & Godfrey, 1985). According to Kruger and Van Breda (2001), the primary goal of EHWPs is to improve the social wellness and operation of the work organisation as a community, by aiding in the accomplishment of the goals of the work organisation and enhancing its efficacy.

Mental healthcare practitioners now have a stronghold due to changes in EHWPs in worksite wellness initiatives. Trained mental healthcare providers can build rapport, demonstrate empathy, and assist service users in reaching their objectives and fulfilling their needs. Drawing upon the

skills and inherent training of mental healthcare professionals reduces the need for additional support and behaviour modification. Due to the services provided by mental healthcare specialists in EHWPs, employers experience more program execution and cost savings (Groenveld *et al.*, 2011; Saliba & Barden, 2017). The development of technology has had an impact on social work and EHWPs, as it has on other scientific fields. Technology appears in various ways to support psychosocial health, including calls, e-mail, chat, video conferencing, and multimedia (Kenny & Mceachern, 2004).

Despite specific concerns about online therapy, this mode of therapy has become a vastly expanding mode of offering services to service users. The first paid online counselling sessions appeared in 1995. The number of sessions available at any given time was not officially documented or determined. However, the rise in the number of therapeutic services provided in this manner demonstrates a clear public need for online counselling (Anthony, 2003; Grover *et al.*, 2002).

According to Hanley and Reynolds (2009), the physical distance between the counsellor and the service user is one of the critical tenets of the traditional therapeutic connection that is broken by Internet therapy. According to critics (Mallen & Vogel, 2005; Schlenger *et al.*, 2022), internet counselling does not achieve sufficient degrees of connection and empathy. Following the COVID-19 pandemic's onset, online therapy modes have become more prominent, which has prompted Sperandeo *et al.* (2021) to investigate whether empathy is taken into account in the same manner in face-to-face and online therapy contexts. The findings proved quite significant in that service users thought their therapists were much more sympathetic and encouraging when they were remote in contrast to in-person sessions. These findings are significant as they documented that positive empathic perception could predict therapeutic outcomes. These outcomes are, however, still the results of human interaction. AI has the limitation of not always being able to read emotions in a text-based format, although advancements in processing text into emotions are encouraging (Wahyono *et al.,* 2019). This process requires algorithms that are based on comment columns in text messages. In this context, AI text mining is required and integrated into the application's learning to identify emotions (Divayana *et al.,* 2017).

The Internet, the IoT, and cloud computing are currently seeing widespread popularisation, where sets of information produced by people and organisations are starting to increase swiftly. In line with Hsu and Hasmath (2017), social workers must confront Internet- and digital thinking directly rather than reacting from afar. The authors believe that a breakthrough must be sought. Cloud computing and mobile computing IT have entirely changed how conventional industries develop. The development of social work services aids the fusion of big data technologies and social work. As big data can capture and update real-time data on human behaviour and respond to it, the future of big data-driven social work will be mirrored in its application. The information will serve

as the foundation for precise and effective technical tools for risk assessment and crisis intervention. Chen and Chen's (2022) core argument for the digitisation of social work is enhancing care and reinforcing professional ethics in social work.

AI and social work seem far apart, however, Tambe and Rice (2018) argue that these disciplines can produce a powerful new science when brought together. The authors argue that Artificial Intelligence for Social Good (AI4SG) - or AI and social work - is a new intellectual space due to collaborations that are unique to both sciences. Chatbots are a promising addition to traditional forms of services, given recent advancements in AI and deep learning (DL), the widespread use of online digital devices and the emergence of vast data sets from digital trails and social media chats (Crutzen, 2011). The potential advantages of chatbots have made this kind of interface an appealing and intriguing way to provide information to mental healthcare users. However, Howick *et al.* (2001) find it hard to imagine that AI carers can empathise, understand and offer the same quality of care humans can. According to Nagarajan and Yuvaraj (2019), additional potential limitations of chatbots include loss of non-verbal cues, the possibility of the client becoming more dependent due to constant contact with a mental healthcare provider, the possibility of the client imitating problems, and concerns surrounding the privacy of information shared online.

This study does not propose replacing mental healthcare professionals with AI technology and chatbot interaction but instead proposes that this product serve a supplementary role in traditional practice. Numerous advantages of online mental health services have been identified, including removing the historically attached stigma from mental healthcare, ease of access and convenience, privacy, comfort and safety, ease of communication, cost- and general effectiveness, improved consistency, more frequent and flexible sessions, better progress, increased access, anonymity, enhanced autonomy, the immediacy of services and motivation to attend in-person services, should it be necessitated (Cherry, 2021; McGee, 2022; Novotney, 2017; Poh Li *et al.*, 2013).

In the South African psychosocial counselling context, organisations such as the host company (Lyra Southern Africa – formerly the Independent Counselling and Advisory Service [ICAS]) in this study, the South African Depression and Anxiety Group, and MobieG have live chat platforms. These services consist of a discussion (or psychosocial counselling session in this instance) between two people that involves sending instant text-based messages over the Internet; these services are employed to supplement traditional telephonic services (Cipoletta *et al.*, 2021). According to Gilat and Shahar (2007), live chat services are particularly feasible for crisis support due to their accessibility, confidentiality, and ease of help-seeking during a crisis. However, the platforms mentioned are human-driven and pose limitations to the service, such as increased personnel costs for 24-hour services; differing engagement- and therapy styles for the same service user; fatigue of counsellors during night/overtime shifts, mistakes due to servicing multiple

clients simultaneously; delayed response times; connectivity challenges of counsellor technology and IT equipment; rigorous administrative obligations; the increasing cost of application, and chat platform development and maintenance. The increased cost of hosting 'direct-messaging software-as-a-service' platforms such as Zendesk and Tidio. Figure 1 provides an example of an existing live chat service and engagement. Please note the time stamps as an indication of the immediacy of the service.

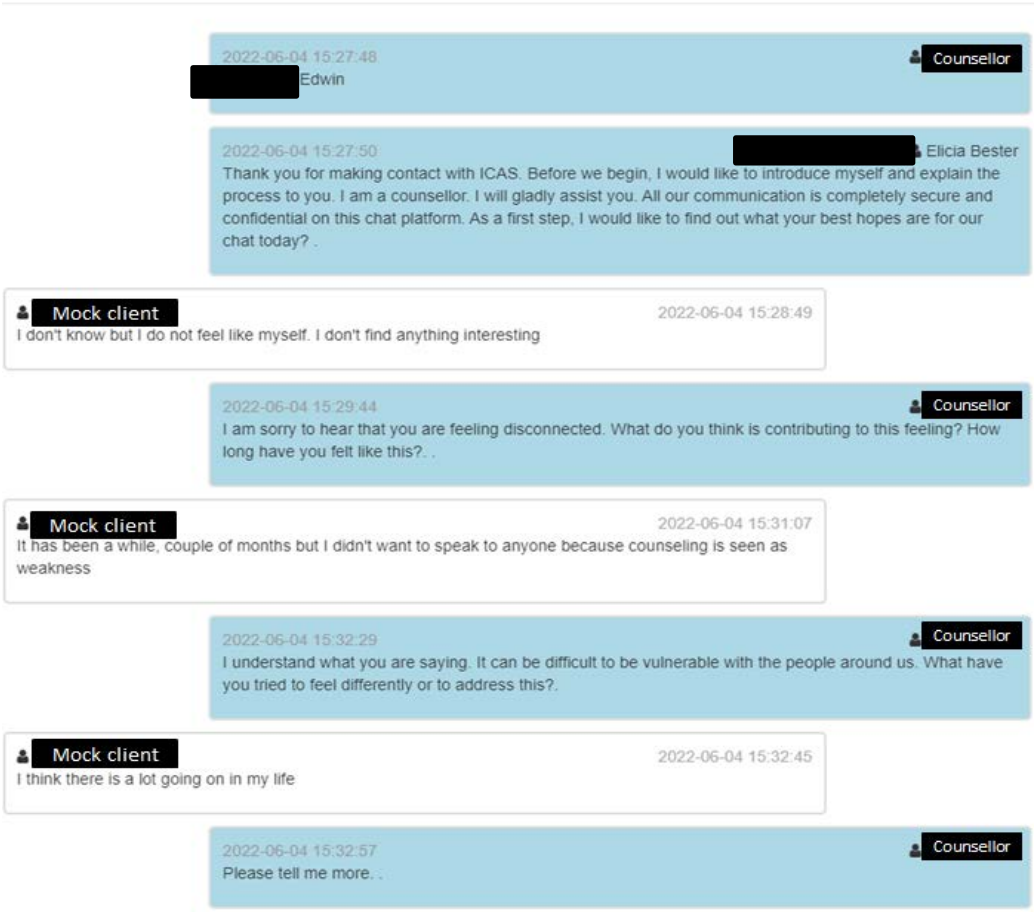


Figure 1. Example of a live chat interaction

The challenges and limitations of live counselling can be addressed by the implementation of AI therapy platforms supplementing existing live chat services. Electronic chatbots ensure that multiple services can be provided simultaneously, as well as that services are more stable as engagement and therapy styles remain uniform and agent fatigue is eliminated. Furthermore, chatbot responses will be immediate, saving considerable time spent on administrative functions. All these benefits depend on whether connectivity challenges can be eliminated and application maintenance can be lowered. These benefits align with Roestenburg's (2021) findings, namely that social workers are ready to respond positively to IT advancements and expect IT advancements to contribute significantly to client-based service levels.

There are, however, no AI therapy platforms available in the field of social work, leaving the discipline lagging behind the 4IR, automation, AI, and the IoT. In the quest to catch up, with digital-led technologies rapidly approaching a tipping point of becoming more widely available and cost-effective (Sunter & Ilbury, 2021), it seems fitting that social work adapts or at least adds to its approach to solving problems. In Skjuve and Brandtzaeg's (2018) view, chatbots may be the answer as a new innovative, interactive way to communicate with mental health users. Based on work by Schuetzlera *et al.* (2020), chatbot conversations can be meaningful and trustworthy.

1.4.2. Problem statement

Given the study's context, it is evident that social work as a profession is lagging in technological development, which could make it outmoded in keeping up with the 4IR's tendencies and trends. The observed problem is that social work does not utilise AI and IoT to contribute toward multi-stakeholder cooperation across academic, social, political, national, and industry boundaries.

Using a multi-phased, mixed-method approach, this study proposes the exploration and design of a conceptual AI generative chatbot counselling framework to inform an AI algorithm, specifically in the field of social work that focuses on the needs of users of a broader EHWP. This study will not achieve the complete development of a chatbot device as this requires multiple large-scale studies; however, it aims to develop the conceptual and reference framework underpinning such a device. Thus, developing the actual device or AI is beyond the scope of this research.

1.5. RESEARCH QUESTION

Based on the contextualisation and problem statement of the study, the following broad research question can be formulated: *What aspects must be considered in developing an AI generative chatbot counselling framework for a social work EHWP (Lyra SA)?*

1.5.1. Research sub-questions

- What literature regarding the use of AI in psychosocial health is available?;
- How do live chat counselling clinicians currently address service user (client) problems through text-based live chat services?;
- What are the textual trends within current live chat services data?;
- What components of the conceptual AI generative chatbot counselling framework inform a multiple-response AI algorithm?; and
- Which inputs from subject matter experts could improve the proposed AI generative chatbot counselling framework?

1.5.2. The expected contribution of the study

This study will contribute to the knowledge base or theory of social work. If successful, the study could contribute to the modernisation of social work, which could lead to social work remaining relevant in terms of 4IR/automation/IoT. Overall, this study expects to benefit the users receiving services from Lyra SA in that they may be enabled to receive services from a chatbot based on the framework developed in this study. This study expects to contribute toward cost savings, relevancy, and an enriched scope of services to clients. Additional opportunities for research in social work counselling technologies may emerge from this study. The researcher recognises that the eventual development of a chatbot depends on a multi-disciplinary collaboration across different disciplines and technological fields, such as the IT field.

1.5.3. Research statement

The development of an AI-generative chatbot counselling framework will lead to the modernisation of the social work knowledge base, the discipline, and the services rendered by social workers. This would promote the future relevancy of the profession within the technologically-enhanced world.

1.5.4. Proposed research product

This study proposes the design and development of a conceptual AI generative chatbot counselling framework to inform a multiple-response AI algorithm. This specifically relates to the field of social work, which focuses on the needs of users of a broader EHWP.

1.6. AIM AND OBJECTIVES

A structured approach to developing an AI generative chatbot framework within a social work and EHWP context is provided by specific aims and objectives. The following section details the overarching aim of the study, as well as a series of focused objectives intended to achieve this aim.

1.6.1. Aim

The study aims to examine and develop a social work conceptual AI generative chatbot counselling framework. This will inform a multiple-response AI algorithm within the context of a broader social work EHWP.

1.6.2. Objectives

The following objectives will achieve the aim of the study:

- **Objective 1:** To conduct a literature study on existing frameworks, technologies and services relying on chatbot technology;

- **Objective 2:** To empirically identify interactive pathways in an empirical study on an existing live-chat-generated textual database. Subsequent to this, Objective 2 aims:
  - **Objective 2.a:** To conduct a qualitative thematic analysis on a sample of live chat transcript data; and
  - **Objective 2.b:** To conduct a quantitative frequency analysis on a complete live chat transcript data set.
- **Objective 3:** To merge and interpret qualitative and quantitative findings (convergence) in a converged set of findings;
- **Objective 4:** To design a conceptual AI generative chatbot counselling framework based on the findings in Objective 3 which informs a multiple-response AI algorithm using the Multinomial Naïve Bayes theorem (Xu *et al.*, 2017) and the Software Development Life Cycle (SDLC); and
- **Objective 5:** To refine the proposed AI generative chatbot counselling framework using the Delphi study method.

1.7. OVERALL RESEARCH APPROACH

This study used a multi-phased mixed-method research approach. According to Rubin and Babbie (2016) and Venkatesh *et al.* (2016), mixed-methods research combines qualitative and quantitative methods in one study to better understand the phenomenon. Furthermore, this integration of the two data types occurs at one or more phases of the research process. As a result, this strategy yields 'meta-inferences' - an integrative perspective of data from both qualitative and quantitative strands - and overcomes the constraints of utilising a single approach in isolation.

The study followed a descriptive design to accurately and methodically describe the features of the phenomenon under investigation. Identifying, evaluating, and synthesising patterns and themes that surfaced throughout the research process allowed for a greater comprehension of the data. The descriptive design facilitated the clear illustration of qualitative themes, and where appropriate, allowed for the incorporation of quantifiable elements, thereby enhancing the mixed-method approach (Doyle *et al*., 2019).

The mixed-method research design had an equal emphasis on qualitative and quantitative data, which followed a qualitative first, quantitative second sequential design (Rubin & Babbie, 2016). According to Terrell (2012), the qualitative first, quantitative second sequential design is the right strategy when designing a framework. Therefore, the study proposed a sequential design in which qualitative themes were first identified from a small sample of the large dataset, whereafter a quantitative study followed on the complete dataset, incorporating the findings from the qualitative study as search codes in the quantitative study (Creswell & Creswell, 2018). Finally, the

qualitative themes were integrated by transferring qualitative themes into quantitative variable codes before the quantitative data were generated. The overall research approach is encapsulated and summarised in Figure 2.

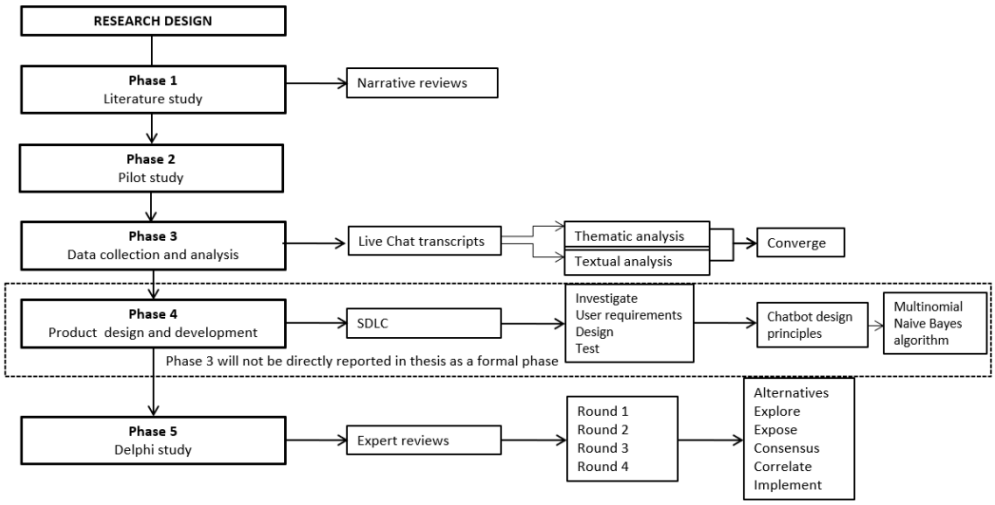


Figure 2. Overall research approach

This research project was conducted through four phases. These include i) a literature study; ii) data collection, analysis and interpretation (qualitative transcript analysis, quantitative textual analysis, and convergence); iii) the design and development of the proposed product, and iv) a Delphi study.

1.7.1. Phase 1: Literature Study

The components of Phase 1 of the study, which includes the description, procedure and data analysis methods are described in the following sub-sections.

1.7.1.1. Description

According to Paul and Criado (2020), it is fitting to commence this project with a literature study, as it will provide an initial exploration and overview in the early chapters, potentially serving as an introductory section to the research. The literature study is intended to study how ideas have evolved and should demonstrate that the researcher has a comprehensive understanding of the existing conceptual framework on the topic (Chalawila & Muchanga, 2022). As per Walliman (2011), the literature study can indicate how ideas are developed and should demonstrate that the researcher has a thorough understanding of the current conceptual framework of the subject.

A literature study is a significant element that acts as a preliminary study and establishes a conceptual and theoretical foundation for a research subject. A literature study enables the researcher to explore, analyse and synthesise published work in a specific domain and draw inferences between existing studies and their own studies (Aveyard, 2010; Shah *et al.*, 2018).

Rubin and Babbie (2016) distinguish between appraising literature reviews as part of a more extensive study of proposals and appraising systematic reviews and meta-analyses, essentially emphasising that the processes are distinct but essential to the research process. The researcher gathered, studied and synthesised literature and meta-analysis to comprehensively understand the existing body of knowledge, thereby identifying shortcomings and establishing a sound conceptual and theoretical framework for the research project.

For this study, the literature study in its simplest sense, is performed as an overview of published materials. It informed two fundamental questions, namely (Munn *et al.*, 2018):

1. What current theoretical issues and arguments surround the topic of chatbots in counselling?; and
2. What is the present state of knowledge on the subject of framework generation?

Additional benefits of the literature study could include, but were not limited to:

- Clarification of key concepts and definitions in the available literature;
- Examination of current research in the field of AI development in counselling; and
- Identification of the concepts' essential traits.

1.7.1.2. Procedure

To produce a high-quality literature study, a modified approach to the process described by Sauer and Seuring (2023) was applied. This involved:

- Step 1: Define the research question;
- Step 2: Determine the studies' required characteristics;
- Step 3: Retrieve a sample of potentially relevant literature;
- Step 4: Select the pertinent literature;
- Step 5: Synthesise the selected literature; and
- Step 6: Report the results.

Once Steps 1-6 were complete, the literature was subjected to critical appraisal criteria as summarised in Figure 3. This figure was adapted from Rubin and Babbie (2013).

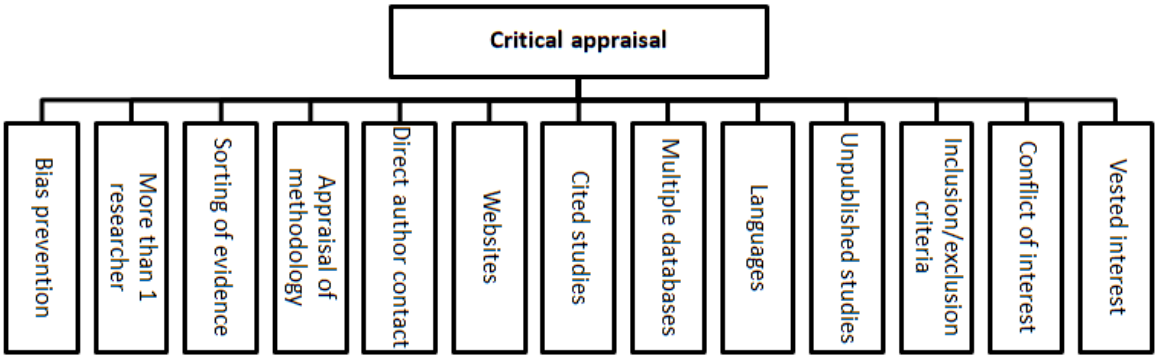


Figure 3. Criteria for appraising systematic reviews and meta-analyses (adapted from Rubin & Babbie, 2013)

The researcher aimed to follow the literature study process, as outlined by Creswell and Creswell (2018). This involved identifying keywords, searching databases for relevant materials, locating the relevant literature, skimming the literature, designing a literature map, drafting summaries of the most pertinent materials, assembling the literature, and reporting on the material.

1.7.1.3. Data analysis methods in literature studies

Rudestam and Newton (1992:49) strongly feel that "your task is to build an argument, not a library." A literature study contextualises the research project by indicating how it fits into a particular field. It aims to identify knowledge gaps, aid in research problem development, identify theoretical frameworks, issues, and variables related to the research topic, and identify conceptual and operational definitions (Kaniki, 2006).

To achieve the objective, the researcher conducted a literature study. This literature study examined new developments in AI applications and counselling rather than just a description or overview of the literature on the topic. The study utilised existing library search engines and the following list of keywords:

- AI;
- AI in mental health;
- 4IR;
- Digital evolution;
- AI in mental health;
- Chatbot(s);
- Human-level ML;
- Superintelligence;
- Naïve Bayes approach;
- Neuro-linguistic programming;
- AI generative chatbot;
- AI4SG;
- IoT;
- Health informatics technologies;
- Virtual mental healthcare;
- Online/Internet therapy/counselling;
- Data-driven social work;
- Digitisation of social work / mental health;
- AI and social work;
- DL;

- ML;
- Transitioning agents;
- Influence maximisation;
- Network structures;
- Mental health prediction models;
- Mental health detection models;
- The history of AI; and
- Future progress in mental health AI.

1.7.2. Phase 2: Pilot Study

The following sub-sections describe Phase 2 of the study, detailing the description, procedure, sampling- and data analysis methods used in the pilot study.

1.7.2.1. Description

A pilot study was conducted, drawing from data collected outside the period of the study sample. The pilot study's purpose was to test the feasibility of the overall study (Strydom, 2021). Additionally, the pilot study tested the data collection methodology, the analysis method and the feasibility of analysing chat-type data.

After the pilot test, the researcher had supervision meetings with the promoters to discuss the outcome of the pilot phase. The meetings concluded that the study was feasible and that the methods rendered significant results. The researcher then proceeded with the main study.

1.7.2.2. Procedure

The following procedure was followed to conduct the pilot study:

1. Collected live chat transcripts from the month before the study;
2. The collected live chat transcripts were anonymised;
3. A research randomiser was used to select 15 random live chat transcripts. The selected live chat transcripts became the primary documents for the pilot study;
4. Primary documents were converted to the correct format for import into ATLAS.ti;
5. Transcripts were studied, whereafter preliminary codes and code families were generated;
6. Accumulative code totals were generated;
7. Codes were linked to code families;
8. A word crunch was completed to determine word frequencies; and
9. Transcripts were exported to ATLAS.ti AI, where sentiment analysis was conducted.

1.7.2.3. Sampling method

The following bullets discuss the sampling method used for the pilot study’s live chat thematic analysis. This includes the sampling (inclusion, exclusion criteria, as well as size and motivation).

- **Live chat thematic analysis - Purposive sampling of total population**

The total population of live chat transcripts during a period before the study was selected for sampling and included in the pilot study. These transcripts were prepared in the prescribed manner and subjected to qualitative thematic analysis and quantitative textual analysis.

- **Sample inclusion criteria**

Live chat transcripts from the relevant period that consisted of 50 or more interactions that were not service enquiries were selected for the pilot study.

- **Sample exclusion criteria**

All live chat service inquiries were excluded from the pilot study as these do not represent counselling sessions.

- **Sample size and motivation**

In total, 15 live chat transcripts from the relevant period were randomly selected and included in the pilot study. These 15 live chat transcripts equated to 16.66% of the data that was used in the study. A 16.66% sample size for inclusion in the pilot study can be justified based on validity, practicality and resource efficiency. By using this method, the sample is guaranteed to be viable and controllable for a pilot study, yet representative enough to yield valuable insights (Giguère-Rancourt *et al.,* 2022).

1.7.2.4. Data analysis

The following sections describe the data analysis for the qualitative live chat transcripts. This includes the qualitative and quantitative analyses.

- **Qualitative analysis**

The transcripts were coded as part of the qualitative study, and themes were found using ATLAS.ti. This procedure made it possible to identify recurrent themes and trends in the live chat sessions.

- **Quantitative analysis**

Word frequency counts were conducted through ATLAS.ti, and sentiment analysis was conducted using ATLAS.ti AI. Insight into the tone and expressions were provided through this analytical method.

1.7.3. Phase 3: Data Collection and Analysis

The researcher collected both qualitative and quantitative data sets sequentially. The qualitative themes were incorporated into the search codes for the quantitative study, to structure the quantitative search.

1.7.3.1. Qualitative live chat transcript collection and analysis

The following bullets describe the qualitative live chat transcript collection and analysis. This includes the description, procedure, and sampling method, size and motivation for the collection and analysis.

- **Description**

The dataset for the qualitative study consisted of a sample of live chat textual transcriptions. These transcripts were the raw data for the qualitative project. Live chat transcripts are verbatim e-mail records of textual chat messages and conversations between clients and counsellors using a psychosocial service rendered by a service provider in the wellness programme sector. The qualitative live chat transcript analysis was conducted by analysing existing live chat text transcripts. The transcript text was, firstly, used as the object of analysis, whereby themes involving problem descriptions, conversational styles, counselling outcomes/performances, grammatical structures, and narratives could be analysed. Secondly, the free-flowing text was used to analyse words and codes using semantic networks and cognitive maps (Bernard, 2005; Guest *et al.*, 2019).

Thematic principles were used in the analysis of the data. Schurink *et al.* (2021) recommend that the researcher first read and reread the transcript text. The aim of this activity is for the researcher to generate provisionally analytic ideas by critically evaluating the meaning of the words. After understanding the data, coding is applied to the data. During this step, the researcher moves to note-taking to systematically code the text through labelling. Pattern identification or theme development follows. This activity consists of identifying themes and patterns that appear frequently, which are achieved by identifying underlying similarities and deviations from the norm. Reviewing (or quality control) of the analytical elements shapes, clarifies, and rejects if necessary. Adjustments to themes and the thematic map are created accordingly. Computer-Assisted Qualitative Data Analysis Software (CAQDAS) or code families constructed of child and parent nodes, were utilised to represent the data analysis.

- **Procedure**

The researcher applied a qualitative thematic data analysis strategy. The following steps were utilised for qualitative live chat transcript data analysis (Braun & Clark, 2006; Creswell, 2014):

1. Preparation, organisation, and familiarising: After collecting the data from live chat transcripts, the researcher became familiar with the data;
2. Reading and generating initial codes: After studying the data several times to get a general understanding, the researcher started to recognise overarching ideas, codes, and patterns in the transcribed data;
3. Searching, reviewing, defining, and naming themes: By regularly evaluating themes and the transcribed data, the researcher, with the help of an assistant coder, named themes and assigned the pertinent data to the various themes;
4. Concept maps: According to Umoquit *et al.* (2013), these maps form a diagram or mind map, which are conceptualised to include the determined themes and sub-themes. Concept maps were applied to the data analysis process to show interconnection and the evolution of a whole. This practice made understanding the connections between concepts and comparing data more accessible; and
5. Data representation and report creation: To assure reliability, the researcher chose substantial extracts from the transcribed data before compiling a reflective report of the studied data.

ATLAS.ti Scientific Software was used to analyse the acquired qualitative data.

- **Sampling method (Live chat thematic analysis - Systematic sampling)**

Live chat transcripts were selected based on inclusion and exclusion criteria. In this instance, chats with more than 50 interactional exchanges that are not service inquiries were included. A tentative population estimate indicated that about 800 live chat transcripts would be available when the study commenced. The population figure is dynamic and changes daily. In total, six sets of 15 cases were randomly selected using a random sampling calculator and by first numbering all eligible cases for participation. Table 1 estimates the number of transcripts based on the number of interaction criteria. Abbott *et al.* (2021:182) refer to this amount of data as “big data”, which is particularly useful for extracting trends, patterns, relationships, and interactions. Table 1, therefore, indicates the possible number of transcripts that could be included when the minimum interaction criteria were raised in increments of 10 over six months.

Table 1. Live chat transcript inclusion criteria

Number of interactions	Number of transcripts for inclusion
30	1611
40	1123
50	782
60	551
70	375
80	276
90	189
100	129
110	85
120	65
130	42
140	27
150	18

Each set of 15 cases was analysed, after which more samples were drawn until preliminary thematic saturation. In addition, Lyra SA provided the live chat transcripts as they are the owners of the data set. The organisation has obtained informed consent from all participants at the start of each engagement in live chat. Chat transcripts for six months were downloaded from Cubit Suite (the organisation's chat software) and served as the population for sampling. The employee wellness organisation anonymised the data first by separating the chat transcript from the identifying details of the service user. The analysis procedure is summarised below:

1. Collect the population of live chat transcripts for inclusion in the sampling frame;
2. Anonymise each transcript by removing any reference to any variable that may identify a participant, followed by downloading;
3. Store the population of transcripts on the researcher's password-protected computer;
3. Number each transcript with a unique identifier;
4. Use a research randomiser to randomly select 90 transcripts, consisting of 15 random transcripts over six months;
5. Generate preliminary codes and nodes using qualitative software such as ATLAS.ti;
6. Generate preliminary family trees consisting of parent and child nodes. These are graphic illustrations of the linkages amongst codes and nodes in the thematic analysis. Taxonomies are critical in the analysis process as they indicate logical linkages amongst themes in the thematic analysis, and these will be used in the design of the counselling App as suggestions for patterns or conglomerates of codes;
7. Randomly select further samples until thematic saturation is achieved; and

8. Once a uniform taxonomy is reached containing an exhaustive code list, a code list is developed from the themes in preparation for the quantitative phase of the study.

For this study, the sample inclusion criteria (for the thematic analysis) included all transcripts with 50 or more interactions between the clinician and the service user. Additionally, the sample exclusion criteria (for the thematic analysis) involved the exclusion of all live chat service inquiries, as these do not represent counselling sessions.

- **Sample size and motivation (Textual data analysis)**

Sampling continued until thematic saturation was achieved. As study designs vary, Guest *et al.* (2006) contend that there is no one-size-fits-all approach to sample saturation. However, researchers concur that thematic saturation has been reached when no new data, themes or codes emerge from the analysis (Fusch & Ness, 2015).

1.7.3.2. Data gathering process

The following bullets describe the textual analysis of the live chat data. This includes informed consent, permissions and the procedure followed in collecting the data.

- **Informed consent from live chat service users**

Informed consent (Figure 3) was obtained from live chat service users through an opt-out option, which was displayed once the client connected to a counsellor before the commencement of the live chat session. An example of the consent message is displayed below (counsellor's view).

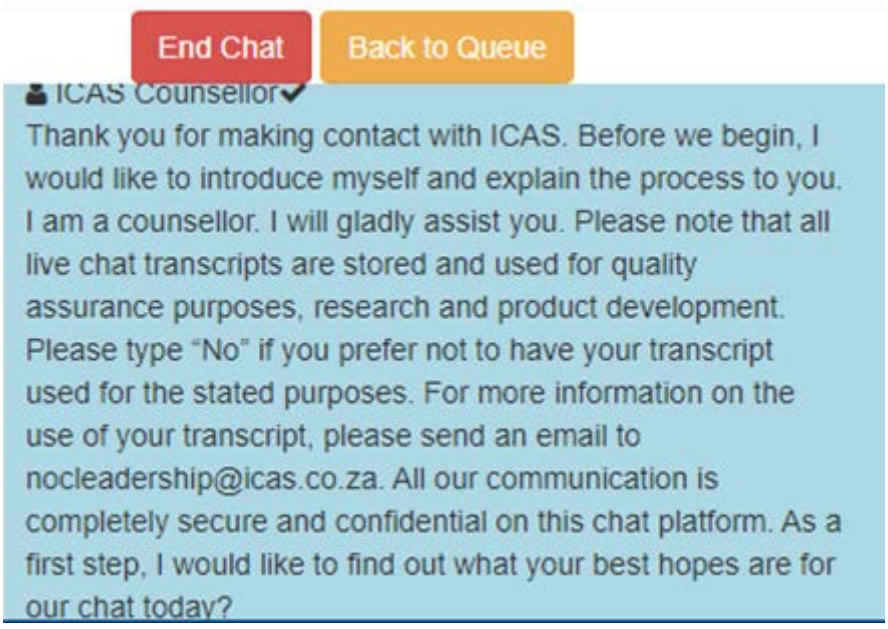


Figure 3. Example of consent (counsellor's view)

The service user's view on their smart device is displayed below (Figure 4). If a client did not consent to the use of their live chat transcript, the transcript was disregarded for possible inclusion in the study.

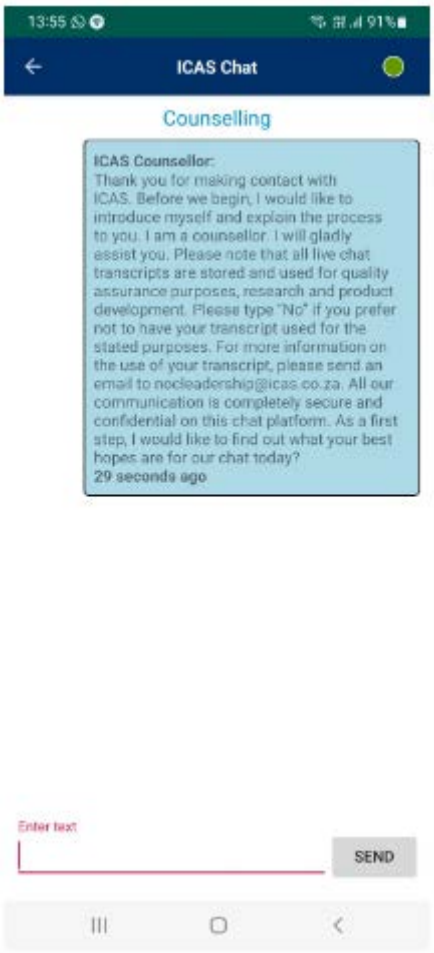


Figure 4. Example of consent (service user view)

- **Permission for the study**

This paper has attached a permission letter to conduct the study within the host company's context. Service users (or the study participants) have consented to use live chat transcripts for research and product improvement purposes before utilising the services.

- **Procedure for obtaining consent from counsellors**

In addition, the researcher sought the consent of the counsellors who participated in the live chats during the data collection period. The researcher sent project advertisements to each counsellor who has worked on the live chat system, inviting them to consent to use their textual data contributions in the live chat system. The flyer explained the project to them, and an opportunity was given for them to ask questions about the project. An independent person conducted the

project information stage. An opportunity was created for counsellors to sign the informed consent, and for this purpose, both the researcher and the counsellor were present in the organisation's office to complete the signing process. Data were only transferred after receiving approval from each counsellor. Live chat transcripts of counsellors who did not consent were removed from the dataset before transferral of the data to the researcher. The company also provided written permission granting the researcher access to the live chat transcripts and they anonymised the data before release.

1.7.3.3. Validity, reliability, and trustworthiness

It is essential to assess the study's validity, applicability, consistency, and objectivity. Thus, there are four additional structures found. These constructs are described by De Vos (2005) as follows:

- How reliable are the study's results?;
- How adaptable and applicable are the results to a different environment or population?;
- If the study was conducted with the same subjects and in the same environment, how can we be confident that the results will be repeated?; and
- How can we be sure the results accurately reflect the participants' opinions and are not the result of the researcher's bias or prejudice?

Creswell (2014) suggests using various techniques that should improve the researcher's capacity to judge the veracity of findings and persuade readers of their veracity. The author lists eight key tactics that can be used to guarantee accurate results. Of those eight, the researcher i) triangulated multiple data sources; ii) offered contradictory or harmful material opposite to the themes found, and iii) described conclusions using rich and detailed language.

The research findings were reliable because data saturation occurred, and comprehensive and detailed data were gathered. It is possible to recode and code the data to verify the accuracy of the obtained data (Botma *et al.*, 2010).

The amalgamation of quantitative and qualitative data improved the value of mixed-methods research (Fetters *et al.*, 2013). By combining the two types of data, several benefits accumulated. For example, it was possible to create a qualitative sample using quantitative data or to utilise quantitative data to explain results from qualitative data. Furthermore, the qualitative data influenced the creation or improvement of the testing of the quantitative findings.

Credibility was ensured using peer debriefing (Anney, 2014). By subjecting themself to probing inquiries from experts prepared to offer academic counsel, the researcher tested the developing insights into the study.

A co-coder was employed to enhance the trustworthiness of the thematic analysis. The co-coder was a person who had a background in research and psychosocial health. First, the researcher

generated initial thematic codes; thereafter, the co-coder confirmed or refuted the generated themes and codes.

Barnes *et al.* (2012) note that transferability is a process by the research readers. They advise readers to take note of the particulars of the study situation and compare them to the particulars of a setting or circumstance they are familiar with. Thick descriptions were gathered, including details and a thorough explanation of the subject studied to ensure transferability.

According to Anney (2014), confirmability is the extent to which other researchers can confirm or corroborate the investigation findings. An audit trail can be provided to establish conformability. Visible proof of the procedure and the outcome can be provided, demonstrating that the researcher did not merely discover what they set out to discover or that the results were the product of their imagination.

1.7.3.4. Quantitative textual data analysis

The following bullets describe the quantitative analysis of live chat transcripts, which used coded terms to quantify and visualise thematic frequencies.

- **Description**

The second sub-phase of the study consisted of the quantitative analysis of live chat transcripts by using the codes as search terms used in analysing the remainder of the population data. The code search is a data collection procedure that renders frequencies of code contents representing densities in sentiments and emotions. The frequencies were quantified and displayed numerically and supplemented using visual graphics. Ray and Chakrabarti (2017) describe this process as ‘lexicon-based sentiment analysis’.

Text analysis involves examining textual material, removing irrelevant information, recognising linguistic patterns, and tokenising the data. Keyword analysis was applied, where keywords were identified and quantified within their context. A visualisation technique, which is the graphical display of extracted features, was also added. Visualisation techniques aid in data's quick and easy interpretation (Fareed, 2021).

According to Catelli *et al.* (2022), a lexicon-based sentiment analysis is a method predicated on the idea that a text's semantic orientation is determined by the polarity of its words and phrases. The approach relies on content- or trigger words, adjectives, adverbs, nouns, and verbs. WordNet is a useful lexicon-based analysis tool for representing sentiment interconnectedness. Figure 5 is an example of text analysis completed in preparation for this paper.



Figure 5. Example of text analysis

- **Procedure**

The researcher aimed to analyse textual data sets by implementing the following procedure, as adapted from Lexalytics (2020):

1. Language identification: All data sets were already in English;
2. Tokenisation: The text was broken into individual units of meaning or tokens;
3. Sentence breaking: Sentence boundaries were determined;
4. Chunking/parsing: The text was broken into component phrases (noun phrases, verb phrases, and so on); and
5. Generate word vectors: The frequency and contexts in which words appear were generated. The frequencies allowed the researcher to prioritise and arrange coded text units in order of density or importance, a critical step in designing the conceptual model in the next phase of the study.

- **Sampling method**

Quantitative textual analysis - Step 9 of systematic sampling in qualitative analysis *(Purposive sampling - Total population sampling)*

The total population of live chat transcripts used for sampling were included in this study. These transcripts were prepared in the prescribed manner and subjected to quantitative textual analysis.

Sample inclusion criteria *(Textual analysis)*

Data sets subjected to a thematic analysis were included for textual analysis as a form of sampling with replacement.

- **Data gathering process**

The coding process for this study was carried out using ATLAS.ti, a qualitative data analysis software designed to facilitate in-depth code exploration and organisation.

- **Data analysis**

The quantitative phase did not require in-depth quantitative analysis and did not require complex statistical consultation. The sequentially collected qualitative and quantitative data were analysed separately, starting with qualitative thematic analysis. After completion of the quantitative analysis, the quantitative results were compared to the qualitative themes to determine if the two sets of results confirmed or disconfirmed one another. Finally, the researcher triangulated the methodologies through a direct comparison of the qualitative and quantitative data, where the results of one phase were independent of the results of another phase, with the goals of corroboration and validation (Demir & Pismek, 2018; Doyle *et al.*, 2016), as indicated in Table 2.

Table 2. Summary of data analysis procedure

Phase	Activity
Qualitative analysis of live chat transcripts	1. Selected six sets of 15 live chat transcripts 2. Analyse thematically 3. Compile a taxonomy of themes
Quantitative analysis of live chat transcripts	1. Total population sampling 2. Conducted a text search using lexicon analysis to get density 3. Showed the frequency of repetitions in codes
Design App specifications	Integrate results from lexicon analysis into a first design.
Delphi Panels to confirm	

- **Textual analysis**

The study used advanced analytic techniques to interpret linguistic textual densities further. The following analysis techniques were employed:

- A data-driven intelligent mining method (Zhang, 2022) using the traditional Bag-of-words model was utilised to conduct this specific textual analysis. Furthermore, to analyse the

textual data frequencies, natural language processing (NLP), an AI-driven text-analysis approach to quantify data contained within a text, was utilised;

- Moore and Ophir (2021) suggest calculating a topic model based on word occurrence using existing text analysis AI, such as MonkeyLearn, RapidMinder, Oracle Data mining or International Business Machines' Statistical Package for the Social Sciences (IBM SPSS) Modeller, that can identify and calculate the thematic content of big data sets;
- Development of a topic model using Latent Dirichlet allocation (a probabilistic generating model for text corpora and other discrete data sets) and Gibbs sampling (Blei *et al.*, 2003) grouped words into topics such as statistical entities, frequency distribution, and common occurrence. Then, an algorithm selected a list of subjects from the data set, modelling each word according to its likelihood of occurrence in the data set; and
- The phrases that are most loaded (prevalence) over each topic were examined to interpret the topics. A theta matrix followed, where a topical network based on associations was calculated. Topics were then converted into nodes, where the size of each node was dependent on the number of links between the nodes (Luther, 2017), and the edges between the topics represented the calculation. Finally, the cosine similarity of the frequency of terms appearing in the same dataset was used to calculate the result (Walter & Ophir, 2019).

The analysis attempted to answer the following questions:

- What were the linguistic styles used by service users?;
- How do chat service users report problems in their initial problem statements?;
- Which therapeutic approaches did live chat counsellors use?;
- Which counselling skills were most likely used by live chat counsellors?;
- What was the linguistic approach of live chat counsellors?; and
- Which approaches/skills/methods appeared to yield positive or negative results?

The above analytical themes were developed in consultation with the study leader and are crucial in determining the "personality" and therapeutic approaches of the proposed AI conceptual framework (Haber *et al.,* 2024). The quantitative data analysis outcomes informed the researchers of the technical factors while developing a conceptual AI generative chatbot counselling framework to inform a multiple-response AI algorithm. Critical aspects from a social work counselling perspective that warranted attention were:

- What assessments need to be completed by an AI;
- What assessment protocols and scales will be used;
- The nature of risk protocols to be followed;
- Routing strategies that need to be considered;

- Alternative routing strategies to be considered;
- Therapeutic approaches and models that would be most beneficial;
- Topics that require psychoeducation to fulfil; and
- Topics that provide further decision-making information.

Following the analysis of Phase 2B data, the supervisor and the researcher conducted the design process in consultation with relevant experts as required. The data informed the proposed framework's 'hard' and 'soft' processes. For this study, the 'hard' processes were the functions (routing, assessment protocols, and so forth) that the proposed product aimed to accomplish, and the 'soft' processes refer to the way (linguistics, questions, among others) in which the goals of the proposed framework were achieved. Finally, a conceptual AI generative chatbot counselling framework to inform a multiple-response AI algorithm was designed based on the arrangement of the functional and content factors as described above.

1.7.3.5. Risks and benefits analysis

Risks were minimal, as per North-West University (NWU) Health Research Ethics Committee's (HREC) risk descriptors. The online chats were analysed after they occurred, thus posing minimal risk as the sessions have already been concluded and are considered retrospective data. Firstly, the chat data was anonymised before analysis, and the identifying details were kept at Lyra SA. Therefore, there was no risk that any person could be identified from the original database. Confidentiality was, therefore, maintained. Furthermore, no reference could made to the names or details of clients. Lastly, there was no manner that a person could identified by analysing textual content, either qualitatively or quantitatively. Therefore, the risks of the study were minimal.

There were no direct benefits to service users and clients, but the study benefitted Lyra SA and the fields of social work and EHWPs, as previously described. The benefits outweigh the risks.

1.7.3.6. Confidentiality, privacy, and anonymity

Since the study did not collect or use observational data obtained from live participants, issues such as privacy, confidentiality and anonymity were entirely enshrined in the project's data management and storage plan.

- **Data management and storage**

All textual data were stored on the lead researcher's password-protected computer in their office and transferred to the NWU Drive for secure storage. Any data analysis was coordinated from

this station. After the project, all data will be transferred to the Community Psychosocial Research (COMPRES) office for secure storage for five years, after which the data will be deleted.

- **Incentives/reimbursement**

This study provided a token of gratitude to the participants of the Delphi panel.

- **Risk level descriptor**

The thematic analysis of the live chat transcripts posed minimal risk to the service users, as the transcripts formed part of the mental health services that the service users have already received. No contact was made with the service users, and no biographical details were obtained - from the NWU Risk level descriptors, note that document analysis carried a 'no to minimal' risk level.

1.7.4. Phase 4: Design and Development of Proposed Product

The third phase involved preparing a design for a conceptual AI generative framework. The researcher used the product development and SDLC framework and meta-paradigm for designing the conceptual framework based on principles of the Multinomial Naïve Bayes algorithm.

1.7.4.1. Description

Based on the data and findings as informed by Phases 1 and 2, the researcher aimed to develop a product/program in the form of a conceptual AI generative chatbot counselling framework to inform an AI algorithm. The product is in the form of a concept proposal. The first document to be finished in the SDLC can be considered a concept proposal. The development of the social work conceptual AI generative chatbot counselling framework is based on the fundamental principles of chatbot design. It should, in all probability, be able to inform the conceptualisation of the Multinomial Naïve Bayes algorithm that could be used in future research and development of an AI counselling agent.

The process of planning and creating a social work program involves innovative problem-solving for a challenging situation. A successful program design would ideally incorporate reliable research knowledge and industry best practices to identify the optimal combination of components for the program's success. Program design and development create new, frequently enhanced services that meet program objectives through an awareness of the problems to be solved. Program design and development aim to develop services with the best results and the greatest likelihood of fulfilling the program's goals to effect positive change (Kettner *et al.*, 2008; Rothman & Thomas, 1994).

According to Goldkind (2022), academics have been debating the integration of AI as expert systems across social work practice since the 1990s. Academics and other experts have long

been discussing deploying neural networks and other predictive models to enhance treatment outcomes in social work interventions (Patterson & Cloud, 1999). Although the concepts underlying the use of AI in social work are not new, the computational power required to create these tools has been insufficient lately. In several fields of social work, AI has already made significant strides. Lisetti *et al.* (2013) point out that, on the individual level, AI is being used to address the discrepancy between mental healthcare users and practitioners available to meet their needs using AI-enhanced interventions. The potential use of technology is often referred to as AI for the social good.

The proposed conceptual counselling framework aimed to inform an AI algorithm that could lead to the development and implementation of an AI generative chatbot. A chatbot powered by AI uses NLP and ML to comprehend the intentions of service users (Fainchtein, 2022). These chatbots can learn and understand language and communicate using natural language at a practically human level. The more circumstances these virtual agents encounter, the more they can develop new communication techniques and develop their intelligence.

There are two types of chatbots that exist, namely closed-domain chatbots and open-domain or generative chatbots. As the name suggests, closed-domain chatbots are confined to a predetermined set of questions and responses. On the other hand, open-domain or generative chatbots can generate multiple responses by drawing upon previously stored data and interactions. If the chatbot does not have the information in its algorithm, it will ask the service user questions, enabling it to learn new and (mostly) relevant responses. To create the proposed concept design, AI forms such as NLP, ML, and semantic understanding will need to be explored.

Krishnav (2002) recommends considering several aspects when designing a chatbot framework. They are discussed in the bullets below.

- **The type of chatbot**

This study aimed to inform the conceptual framework for the algorithm of a text-based self-learning generative chatbot. Furthermore, the study aimed to inform the creation of a so-called future chatbot that can communicate on several levels with automation at the system level, which refers to the chatbot's ability to maintain the system, task, and personal context.

- **Architecture**

A DL model based on NLP, corpus or data training, a chat window, session or front-end application interface, as well as an application database for processing activities to be performed, should make up the conventional architecture of the chatbot. This study aimed to inform a Multinomial Naïve Bayes algorithm, a data classification technique and an inductive research

procedure for the most effective and efficient ML (Susanti *et al.*, 2017), as well as the NLP process. The corpus training was proposed based on the live chat transcript analysis. The architectural model is displayed in Figure 6.

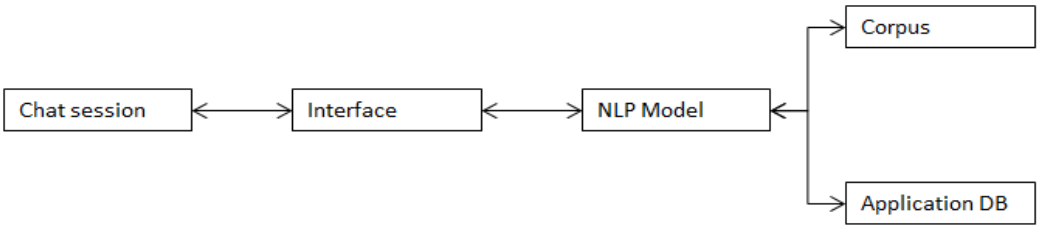


Figure 6. The architecture of a typical chatbot

- **Corpus or data training**

The data on which the NLP model could be trained is called the corpus. This study used regular expressions yielded by the quantitative textual analysis as the corpus training data set. According to Kirsten (2022), non-relational unstructured data tables should sufficiently inform object properties and values.

- **Algorithm**

The corpus training model will inform the algorithm, which should lead to the design of a natural language toolkit (NLTK). This could eventually be utilised in developing a neural network in future multi-disciplinary research.

The concept design aims not to replace human practitioners but to design a generative AI agent framework that could assist live chat counsellors, telephone counsellors, e-mail counsellors, and virtual counsellors. The AI agent could assist in various ways, such as (Beg *et al.,* 2024; Potts *et al.,* 2021):

- Service users can engage in brief conversations;
- Immediate feedback to the service user;
- Completion of various assessments, which could then be routed to alternative or additional services;
- The agent could serve as a catchment for service users who need and might not receive immediate services;
- Information on alternative or additional services could be provided;
- Psychoeducation could be provided to the service user;
- Assistance to the service user with decision-making;
- Assistance with therapeutic homework tracking could be provided;

- Mood tracking of the service user's mood states;
- Evidence-based lessons could be provided to the service user;
- Empathic validation and responses to service user feelings;
- References to previous conversations;
- The service user could request additional services, such as a telephone call by a counsellor, e-mail counselling, and so on;
- Predictive routing to additional/alternative services could be made; and
- Service user experience could ultimately be increased.

The aforementioned supplementary services could be achieved by using the proposed product. The proposed product conceptual framework process flow, as illustrated in Figure 7, would conceptually involve a sequential process that would commence with the client connecting to the platform via a mobile application. An opening statement by the AI agent will be made, and the service user's biographical details will be obtained if not previously obtained. The service user will then respond with a statement. This statement will route the client to an assessment or a follow-up client-initiated therapy session. If the client is routed to an assessment, the assessment will be based on the client's initial or problem statement. The study proposes that scientifically sound assessment scales be utilised to complete the assessments. For example, if the client's initial statement indicates anxiety as a primary problem, the client will be prompted to complete the Pennsylvania State Worry Questionnaire (PSWQ). Automatic calculation of the responses to the assessment provides feedback to the client. If a risk (such as suicidal ideation) is disclosed, the National Support Centre (operations centre of Lyra SA) will be informed electronically. A human clinician can then manage any risks reported in real-time. After the assessment and feedback to the service user, the service user will have the option to commence with a therapeutic process or have the option to be routed to a human clinician or alternative services/resources.

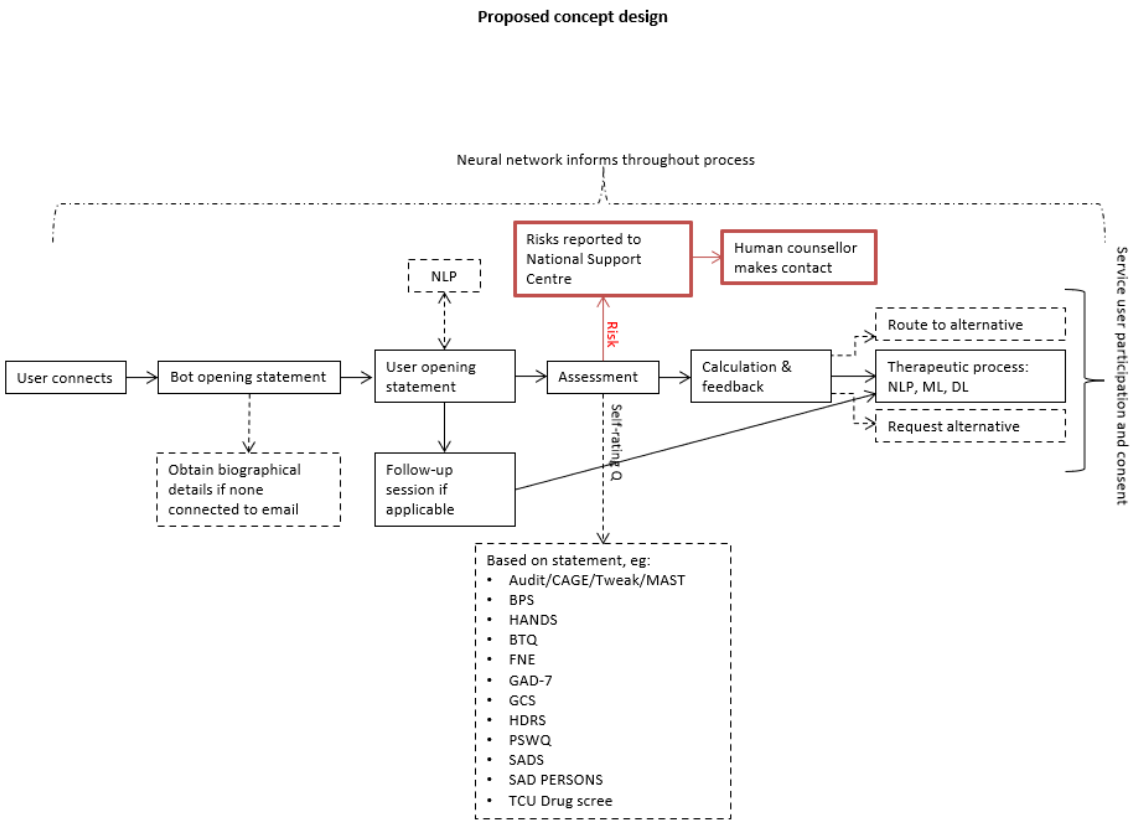


Figure 7. Proposed product concept design/process flow

1.7.4.2. Procedure

There are various descriptions of the SDLC. The descriptions differ from field to field for which a system has been developed. The life cycle is a plan to develop a quality, usable product optimally. The plan, in turn, is not a strict step-by-step procedure but a flexible approach that guarantees consumers, designers, and managers participate in the creation of the final product directly. A typical life cycle for a concept such as this would consist of six stages, namely (Hoffer, 2002; Jirava, 2004):

- 1. Investigation;
- 2. User requirements;
- 3. Analysis;
- 4. Design;
- 5. Testing; and
- 6. Implementation and maintenance.

The waterfall SDLC model, a sequential software development process (Bassil, 2012), informed Phase 4 of this study. The SDLC model's progress flows downwards through a list of phrases which need to be completed to develop the proposed product. Phases 1-4 of the waterfall model

within the SDLC, adapted from Pinciroli *et al.* (2022) and schematically displayed in Figure 8, were completed to design the concept product.

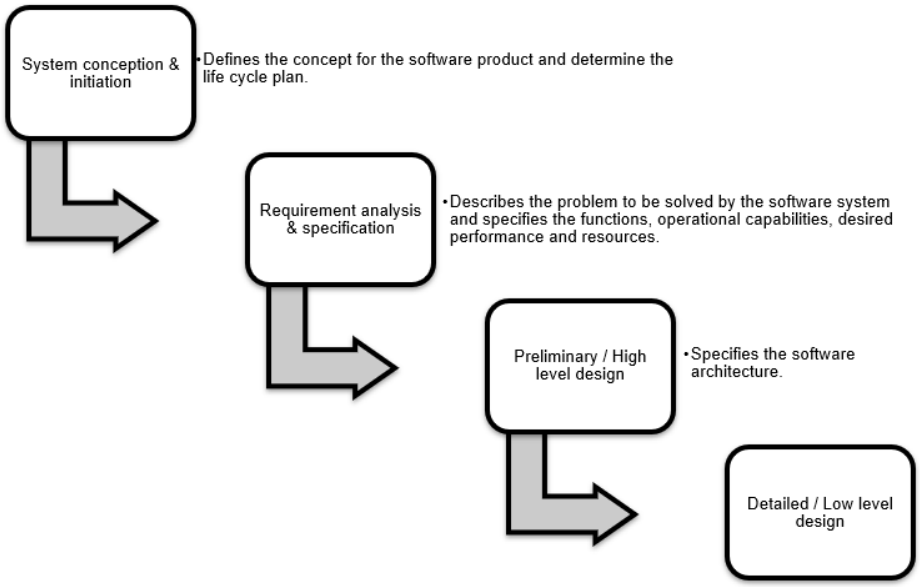


Figure 8. Waterfall model of the SDLC

This study did not aim to implement and maintain (Phase 6) the proposed product, as this would require multi-disciplinary collaboration, a substantial budget and time, which is beyond this study's scope and time frame. Instead, this study aimed to design the concept model for the AI agent, which can be implemented and maintained through further studies and resources. The basic DL generative chatbot design (Barker, 2021) that was followed included:

1. The preparation of data;
2. Data reshaping;
3. Pre-processing;
4. Chatbot type selection;
5. Generation of word vectors; and
6. Seq2Seq model creation.

Further steps within the sequence of chatbot design were not followed, as this would have superseded the scope of this study and required various multi-disciplinary resources. With or without assistance from relevant consultants, the researcher and their supervisors carried out most of the activities outlined for the framework design phase. The design was prepared for submission to a Delphi panel in Phase 4 of the study.

1.7.5. Phase 5: Delphi Study

Phase 5 employed the Delphi technique to refine and validate the proposed framework through expert consensus. The following sub-sections describe the Delphi study’s description, procedure,

sampling method, sample size and sample rationale, detailing each aspect's contribution to the refinement of the framework.

1.7.5.1. Description

The study's final phase entailed refining the proposed product by utilising the Delphi technique. According to de Meyrick (2003), the Delphi technique is well suited to research informing health promotion studies. In principle, the Delphi technique is a group participation technique utilised by a researcher, whereby a panel of subject matter experts is assembled, who are then posed with questions and synthesis feedback, guiding the group towards agreement (Donohoe *et al.*, 2012). According to Sullivan (2011) and Hsu and Sanford (2007), the Delphi technique is a strategy that calls upon and analyses the collective subject matter expert judgements to produce results superior to that of a single subject matter expert, which in turn results in increased content validity. Therefore, the Delphi method should, ideally, contain the use of subject matter experts within a panel who remain anonymous and participate in a series of rounds consisting of iterations and feedback (Kasi *et al.*, 2008; Rowe & Wright, 2005).

1.7.5.2 Procedure

By following the procedure above, the following objectives can be achieved:

- Determine if there are product alternatives;
- Expose and explore underlying assumptions that lead to different judgements;
- Seek information that could generate a consensus amongst the experts; and
- Correlate information on the proposed framework emanating from different experts.

Linstone and Turoff (1975) believe that three to four rounds are often adequate to achieve response stability. However, the repetition and minimal variation in subsequent rounds appear to be a trend that proves to be problematic for the participants.

The researcher aimed to implement four rounds, as suggested by Skinner *et al.* (2015). The rounds consisted of and aimed to achieve the following:

- **Round 1:** Explore the subject under investigation. Each participant contributes additional pertinent information;
- **Round 2:** Reach consensus on how the group understands the proposed product and its functioning with additional recommendations;
- **Round 3:** Explore disagreements, agreements, and recommendations – identify the underlying reasons and evaluate; and
- **Round 4:** Final evaluation, adaptation, and feedback to participants via a Delphi panel discussion group.

The Delphi panel met virtually in the form of a Delphi Panel discussion group (e-Delphi) via Microsoft Teams to discuss their findings and inputs. According to Brüggen and Willems (2009), the e-Delphi technique requires less logistical preparation, which makes it more efficient in terms of both time as well as funding. It offers an approach to developing consensus through a methodical procedure. Due to the methodologies employed by e-Delphi and the online discussion groups that were structured, the data quality was more dependable. Online discussion groups exhibited a modest level of involvement, however because of their controlled methodology, e-Delphi had a lower level of engagement but greater completion rates. Mead and Mosely (2001) further state that the combination of discussion groups with the Delphi method is sufficient in data collection, builds structured consensus, minimises bias and groupthink and enhances participant engagement.

1.7.5.3. Population

The population applied to the experts requested to participate in Phase 4 (Delphi study) of the research project. The expert panel for this study consisted of experts from the psychosocial health and IT fields. Social workers, psychologists, psychological counsellors, application developers, and information and communication technology (ICT) experts were requested to analyse, review, and comment on the proposed product. The sample was selected based on the expertise of prospective participants in the fields of psychosocial health, as well as ICT sampling - which was in the form of purposive sampling - where the population was represented by the sample's typical or relevant characteristics (Strydom, 2021).

The study required expert feedback to inform, refine, and validate the designed AI counselling framework through a Delphi panel as set out in research protocol. Prospective participants for the Delphi panel were recruited through an advertisement distributed to candidates within the Lyra SA affiliate network who met the inclusion criteria, as identified by the Lyra SA clinical manager. To maintain objectivity, the recruitment process was conducted by an independent consultant, contracted with Lyra SA, rather than by the researcher to avoid potential conflicts of interest.

The informed consent form was sent to eligible candidates, and initial contact occurred via email. Once a candidate expressed interest, further correspondence proceeded between the consultant and the candidate. The consultant, with details provided by Lyra SA's clinical manager, allowed candidates to ask questions and clarify expectations about the project. A Microsoft Teams meeting was then scheduled for an introductory session. During the introductory session, candidates completed and signed the informed consent form with the independent consultant present, ensuring procedural integrity. The researcher signed the form during this same meeting, and both parties' signatures were obtained via file transfer. To safeguard confidentiality,

participants' cameras were kept off, and identifiable names were removed, ensuring anonymity throughout the sessions.

The sessions were designed with minimal risk, as discussions centred on academic, and intellectual input related to the AI framework, thus reducing any likelihood of emotional or health-related disclosures. Therefore, no specific precautionary steps were required. Participants were informed of the confidentiality limitations in group discussions, and confidentiality rules were established. Data management followed strict protocols, with recordings deleted from Teams post-session, and securely stored on the researcher's password-protected computer.

The sample inclusion criteria were based on participants' expertise in psychosocial health, and ICT. Psychosocial health and ICT participants were identified by the Lyra SA clinical manager based on their qualifications and experience. Data analysis was performed as stipulated in the research protocol, and results were incorporated into a refined version of the framework.

1.7.5.4. Sampling method

Purposive sampling is proposed when the researcher wants to understand situations better. Predetermined selection criteria were implemented as a means of purposive sampling, guiding the sampling process. Purposive sampling can be used in research in various contexts, including preliminary studies where the researcher is still determining whether a proposed study is feasible, case studies, sampling informants with a particular type of knowledge or skill, and when the population is too small for a random sample (Babbie, 2017; Li *et al.,* 2006; Tongco, 2007).

In this study, the researcher purposively selected experts based on their known expertise in the sciences of psychosocial health and ICT as the main inclusion criterion. The level of expertise was assessed by considering the professional qualifications, years of experience, and positions held by prospective panel members. Experts were required to have practical experience in psychosocial health or ICT, relevant academic or professional qualifications, and current or prior engagement in fields related to mental health, AI, or digital health technologies.

The clinical manager at Lyra SA, who was familiar with the backgrounds and expertise of candidates, assisted in identifying suitable experts. Prospective panel members' expertise was judged based on their professional contributions, professional affiliations, or recognisable achievements in their respective fields. Only those meeting these criteria were invited to participate, ensuring the panel was comprised of individuals with a high level of expertise relevant to the study's focus.

- **Sample inclusion criteria**

Experts in the psychosocial health sciences and ICT experts were included in the study. The researcher aimed to recruit experts from various domains, such as academics, psychosocial experts employed by Lyra SA (as well as not employed by Lyra SA), independent practitioners/contractors and authors of relevant material (journal articles, books and videos, among others).

1.7.5.5. Sample size and motivation

According to Ogbeifun *et al.* (2016), a Delphi panel may be as small as three members or as large as 80. This study included 10 experts, as more than 10 would pose the potential challenge that experts influence each other's ideas (Hasson *et al.*, 2000).

1.7.5.6. Data gathering process

The following bullet describes the process that was followed to gain access to Delphi-panellists.

- **Access to participants and data**

Prospective Delphi-panel participants were recruited by distributing an advertisement to possible candidates, which could be identified by the Lyra SA clinical manager, who acted as a mediator. The Lyra SA clinical manager was responsible for identifying potential participants based on the inclusion criteria outlined in the research proposal and leveraging their knowledge of the candidates' qualifications and roles within the Lyra SA affiliate network and the broader psychosocial and ICT fields. The informed consent form was sent to these candidates, and they were contacted by an independent person who allowed them to ask questions about the project and its expectations. After that, a Microsoft Teams meeting date and time were set to discuss the project and its expectations. Participants were then allowed to complete and sign the informed consent form in the presence of an independent person, and a file transfer was used to obtain the signed document from the researcher, who signed it in the Microsoft Teams meeting along with a witness.

1.7.5.7. Validity, reliability, and trustworthiness

Delphi studies are based on a consensus approach, which ensures trustworthiness. Hasson and Keeney (2011) cite the methodological trinity of reliability, validity, and trustworthiness as characteristics of Delphi studies. Continuous improvement and panellist comments are used to reach this degree of trustworthiness.

1.7.5.8. Data analysis methods

Data analysis of the Delphi study was, initially, qualitative. Participants were requested to comment on the proposed product design in adherence to the SDLC and the recommendation of a conceptual framework for a multiple-response AI chatbot algorithm. Comments were regarded as qualitative responses. The analysis consisted of consideration of each comment's merit and consideration of whether alternatives in the design needed to be implemented, as per the researcher in consultation with the supervisors. When the final round of the Delphi was warranted, the researcher no longer provided qualitative, open-ended questions with each item but only entered a quantitative question with each item requesting the expert to indicate whether the design meets their specific criteria. An overall agreement percentage of 70 indicated the Delphi process's finalisation.

1.7.5.9. Risk level descriptors

The Delphi study posed a minimal, low or negligible risk factor, as the anticipated potential for harm or discomfort was no more significant in and of itself than those encountered in the participant's daily life and occupational setting. The risk descriptors of the NWU give an example of research focusing on opinions and perceptions. During the fourth phase of the study, participants were requested to provide their professional input about the proposed product. Minimal risk was indicated, as the participants were only consulted about their specific field of expertise. Participants' contributions were relevant to the design and development of the proposed product. The Delphi expert panellists were expected to provide information on professional opinion and not personal information. A code only known to the researcher was assigned to the participants. Participants' privacy was protected as their information was kept anonymous and not be traceable.

1.7.5.10. Risks and benefits

Risks were minimal as per HREC risk descriptors. Delphi panel members were not exposed to any information that has the potential to inflict harm. Delphi panel members were only requested to evaluate and provide expert inputs on the proposed product's design. The Delphi panellists did not benefit from participating. The researcher benefits from the Delphi panels by receiving feedback on the derived conceptual framework, which goes towards providing a basis for developing future AI-based technologies.

1.7.5.11. Further ethical considerations

The following bullets discuss the ethical considerations pertaining to the Delphi study. These considerations include incentives/reimbursement, confidentiality, privacy, anonymity, the storage of data and the archiving of data.

- **Incentives/reimbursement**

A Takealot gift voucher to the value of R50 was sent to the participants to compensate for time spent/assistance with the Delphi study. No reimbursement was offered to the live chat counsellors, as only retrospective data was used. No additional input is required and no inconveniences were placed on the live chat counsellors.

- **Confidentiality, privacy, and anonymity**

Delphi panellists worked in their own time, in the privacy of their homes and whenever they had time. Therefore, their contributions may be known to the researcher but not necessarily to the other members of the panel. However, should other panellists know of their details, this will not harm them but rather lead to professional admiration.

- **Storage and archiving of data**

Delphi panels do not generate direct, interactive data *per se*. Instead, panellists made their comments on a document, in this instance, the design of the conceptual framework. Their comments were considered and used to improve, expand, edit or reduce the base document, converting it into a workable and acceptable output that served as the intellectual contribution of this study. Versions of the un-refined and refined documents with comments were stored on the researcher's password-protected computer as evidence of the quality enhancements contributed by the panellists. These documents will be stored for five years at COMPRES before being deleted.

1.8. RESEARCH BUDGET

The allocation of research funds is delineated in Table 3. The researcher has assumed responsibility for the associated expenditures.

Table 3. Research budget

Item	Amount [R]
Administrative costs	2000
Binding	870
Language and technical review	12000

Takealot vouchers	500
Total	**15370**

1.9. TIMEFRAME

The project has been planned according to a very stringent timeline. This timeline is shown in Figure 9 below.

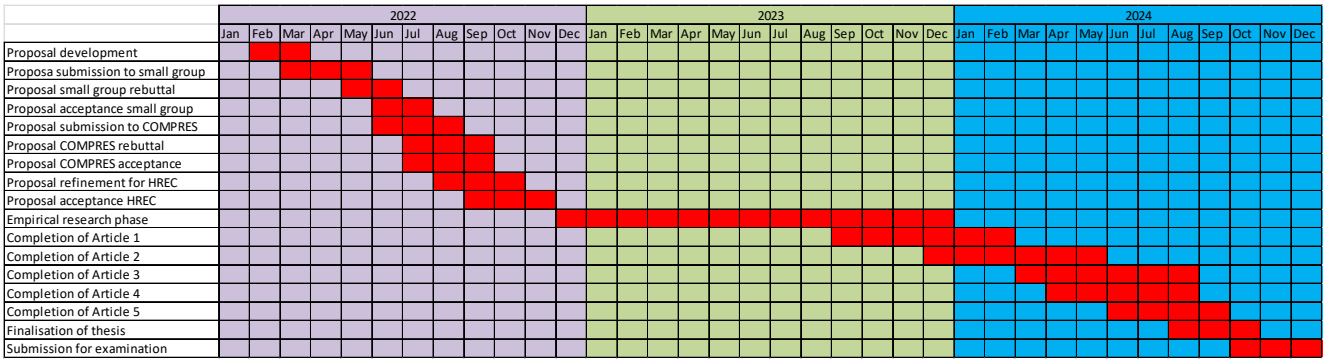


Figure 9. Gantt Chart of timeframe

1.10. CHOICE AND STRUCTURE OF REPORT

The research report is presented in multiple formats. The structure is as follows:

- Section A: Foundations of the Research and Methodological Framework:
  - Chapter 1: Research methodology;
  - Chapter 2: Literature study;
  - Chapter 3: Qualitative transcript analysis, quantitative text analysis and data integration;
  - Chapter 4: Design and development of the proposed product including the Delphi panel; and
  - Chapter 5: Findings summary and recommendations.
- Articles, written according to the editorial guidelines of the journals to which they were submitted for review and publication (not contained in this manuscript):
  - Manuscript 1: “Developing an intelligent online counselling framework for an EHWP: a literature review” submitted for review and publication in *Social Work/Maatskaplike Werk*;
  - Manuscript 2: “Transforming Mental Health Support: AI Chatbots and Insights from Live Chat” submitted for review and publication *International Social Work*; and
  - Manuscript 3: “An intelligent online counselling framework for an EHWP” submitted for review and publication *International Social Work.*
- Section B: Bibliography as guided by the NWU Harvard system.

1.11. EXPERTISE OF RESEARCHERS

The researcher is qualified to carry out the study as they are employed by Lyra SA and have working knowledge of the live chat system in use at this organisation. They have also completed a Master's degree and are familiar with research methods. In addition, they are registered as a social worker with the South African Council for Social Service Professions (SACSSP) and have previously studied under the supervision of one of the current study leaders.

The researcher has received training at NWU, Potchefstroom Campus, predominantly in gathering and analysing quantitative and qualitative data. In addition, the researcher has previously undergone training regarding the basics of Health Ethics research provided by the HREC. The researcher will, however, need ongoing support and monitoring by the supervisors.

1.12. RISK LEVEL OF THE STUDY

The study carries minimal risk since the conceptual framework developed in the study is only an exploratory framework underpinning future technological development. No relationship exists between this research output and any technological or medical device. It is also expected that any future device developed using this framework will be properly and extensively tested before use.

1.13. CONFLICT OF INTEREST

The researcher is an employee at Lyra SA, the company approached for this project. Their employment situation in itself may constitute a conflict of interest; however, apart from having a working knowledge of the live chat system, they have no relationship with any of the participants as none of these were their clients. Any conflict of interest emanating from transactions regarding data transfer will be managed by the supervisors who do not have any affiliation with the organisation.

1.14. DISSEMINATION OF RESULTS

The study's results have been compiled into a thesis, and articles were submitted to recognised social work publications. The researcher communicated the study's findings by publishing the results in a short report that was published on the organisation's website. Any live chat users interested in the study results can click a link to view the findings. Other researchers can utilise the findings to further their research by posting the results.

REFERENCES

Abbott, S., Ade-Ibijola, A. & Roestenburg.W.J.H. 2021. Big data in social science research. In: Fouche', C.B., Strydom, H. & Roestenburg, W.J.H. eds. *Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 181-198.

Anney, V.N. 2014. Ensuring the quality of the findings of qualitative research: looking at trustworthiness criteria. *Journal of emerging trends in educational research and policy studies*, 5(2):272-281.

Anthony, K. 2003. The use and role of technology in counselling and psychotherapy. In: Goss, S. & Anthony, K., eds. *Technology in counselling and psychotherapy: A practitioner's guide*. London: Palgrave. pp. 13-34.

Ausman, M.C. 2019. Artificial Intelligence's Impact on Mental Health Treatments. Paper delivered at the AAAI/ACM Conference on AI, Ethics, and Society (AIES' 19), Honolulu. https://doi.org/10.1145/3306618.3314311 Date of access: 10 Mar. 2022.

Aveyard, H. 2010. *Doing a literature review in health and social care: A practical guide.* 2nd ed. Berkshire, Great Britain: Open University Press.

Babbie, E.R. 2017. The Basics of Social Research. 7th ed. Boston: Cengage Learning.

Barker, S. 2021. Deep Learning Chatbot: Everything You Need to Know. https://shanebarker.com/blog/deep-learning-chatbot/ Date of access: 30 Apr. 2022.

Barnes, J., Conrad, K., Demont-Heinrich, C., Graziano, M., Kowalski, D., Neufeld, J., Zamora, J. & Palmquist, M. 2012. *Generalizability and transferability*. http://writing.colostate.edu/guides/guide.cfm?guideid=65 Date of access: 11 Mar. 2022.

Bassil, Y. 2012. A Simulation Model for the Waterfall Software Development Life Cycle. *International Journal of Engineering & Technology (iJET)*, 2(5):2049-3444.

Beg, M.J., Verma, M., M, V.C.K.M. & Verma, M.K. 2024. Artificial Intelligence for Psychotherapy: A review of the current state and future directions. *Indian Journal of Psychological Medicine.* (July, 28).

Benyera, E. 2021. The Fourth Industrial Revolution and the Recolonisation of Africa: The Coloniality of Data. New York: Routledge.

Bernard, H. R. 1996. *Qualitative data, quantitative analysis*. The Cultural Anthropology Methods Journal, 8(1):9–11.

Blei, D.M., Ng, A.Y. & Jordan, M.I. 2003. Latent Dirichl*et al*location. *Journal of Machine Learning Research*, 3:993-1022.

Boslaugh, S. 2007. *Secondary Data Sources for Public Health: A Practical Guide*. St Louis: Washington University.

Brüggen, E. & Willems, P., 2009. A critical comparison of offline focus groups, online focus groups and e-Delphi. *International Journal of Market Research*, 51(3):1-15.

Botma, Y., Greeff, M., Mulaudzi, F.M. & Wright, S.C.D. 2010. *Research in health science*. Cape Town: Heinemann, Pearson Education South Africa.

Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology,* 3:77-101.

Catelli, R, Pelosi, S. & Esposito, M. Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. *Electronics,* 11, art. #374. doi: 10.3390/electronics11030374

Chalawila, I. & Muchanga, M., 2022. Challenges experienced by postgraduate candidates in the application of conceptual frameworks in scientific Research. *International Journal of Scientific Research and Management (IJSRM),* 10(2):2174-2183.

Chen, Z. & Chen, J. 2022. Decision-Making Model of Social Work Specialization Process Using Big Data Analysis and Chaos Cloud. *Wireless Communications and Mobile Computing*, 1-11.

Cherry, K. 2021. *The Pros and Cons of Online Therapy.* https://www.verywellmind.com/advantages-and-disadvantages-of-online-therapy-2795225#:~:text=Online%20access%20makes%20it%20easier,help%20you%20become%20psychologically%20stronger Date of access: 06 Apr. 2022.

Cipoletta, S., Entilli, L., Bettio, F. & De Leo, D. 2021. Live-Chat Support for People Bereaved by Suicide: A Thematic Analysis. *Crisis,* 43(2):98-104.

Creswell, J.W. & Creswell, J.D. 2018. *Research design: qualitative, quantitative and mixed methods approach.* 5th ed. London: Sage Publications.

Creswell, J.W. 2014. *Research design: qualitative, quantitative and mixed methods approach.* London: Sage Publications.

Crutzen, R. 2011. An Artificially Intelligent Chat Agent that Answers Adolescents' Questions Related to Sex, Drugs and Alcohol: An Exploratory Study. *Journal of Adolescent Health,* 48(5): 514-19.

de Meyrick, J. 2003. The Delphi method and health research. *Health Education*, 103(1):7-16.

De Jesus, A. 2019. *Chatbots for Mental Health and Therapy – Comparing 5 Current Apps and Use Cases.* https://emerj.com/ai-application-comparisons/Chatbots-mental-health-therapy-comparing-5-current-apps-use-cases/ Date of access: 10 Mar. 2022.

Deloitte Development. 2018. *The Fourth Industrial Revolution is here—are you ready?* Deloitte Insights.

Demir, S.B. & Pismek, N. 2018. A Convergent Parallel Mixed-Methods Study of Controversial Issues in Social Studies Classes: A Clash of Ideologies. *Educational sciences: theory & practice*, 18(1):119-149.

De Vos, A.S. 2005. Qualitative data analysis and interpretation. In: De Vos, A.S., Strydom, H., Fouché, C.B. & Delport, C.S.L. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 333-354.

Divayana, D.G.H., Sappaile, B.I., Pujawan, I., Dibia, I.K., Artaningsih, L., Sundayana, I. & Sugiharni, G.A.D. 2017. An evaluation of the instructional process of the expert system course program by using the mobile technology-based CSE-UCLA model*. International Journal of Interactive Mobile Technologies*, 11(6):18-31.

Donohoe, H., Stellefson, M. & Tennant, B. 2012. Advantages and Limitations of the e-Delphi Technique: Implications for Health Education Researchers. *American Journal of Health Education*, 43(1):38-46.

Doyle, L., McCabe, C., Keogh, B., Brady, A. & McCann, M. 2019. An overview of the qualitative descriptive design within nursing research. *Journal of Research in Nursing*. 25(5):443-455.

Doyle, L., Brandy, A. & Byrne, G. 2016. An overview of mixed methods research – revisited. *Journal of Research in Nursing,* 21(8):623-635.

Fainchtein, L. 2020. 28 June. *Generative vs Retrieval Based Chatbots: A Quick Guide* [Blog post]. https://blog.cloudboost.io/generative-vs-retrieval-based-chatbots-a-quick-guide-8d19edb1d645 Date of access: 01 May. 2022.

Fareed, S.B.S. 2021. *API Pipeline for Visualising Text Analytics Features of Twitter Texts.* Paper delivered at the 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), Al Hawiyah, Saudi Arabia. https://ieeexplore.ieee.org/document/9430213 Date of access: 30 Apr. 2022.

Fetters, M.D., Curry, L.A. & Creswell, J.W. 2013. Achieving Integration in Mixed Methods Designs—Principles and Practices. *Health Services Research*, 48(6):2134-2156.

Folstad, A., Brandtzaeg, P.B., Feltwell, T., Effie, L.C., Tscheligi, M. & Luger, E.A. 2018. *Chatbots for Social Good.* Paper delivered at the C.H.I. Conference on Human Factors in Computing Systems, Montreal. https://sintef.brage.unit.no/sintef-xmlui/bitstream/handle/11250/2500523/CHI18_SIG_Chatbots%2Bfor%2Bsocial%2Bgood_authors%2Bversion.pdf?sequence=4 Date of access: 17 Mar. 2022.

Fouché, C.B. & Strydom, H. 2021. Mixed methods research. In: Fouche', C.B., Stryfom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 419-459.

Fusch, P.I. and Ness, L.R. 2015. Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report,* 20(9): 1408-1416.

Gilat, I., & Shahar, G. 2007. Emotional first aid for a suicide crisis: comparison between Telephonic hotline and internet. *Psychiatry: Interpersonal and Biological Processes*, 70(1):12-18.

Glen, S. 2022. *What are Probability and Statistics?* https://www.statisticshowto.com/ Date of access: 18 Mar. 2022.

Goldkind, L. 2022. Social Work and Artificial Intelligence: Into the Matrix. *Social Work* (Provisionally Accepted). doi: 10.1093/sw/swab028

Googins, B. & Godfrey, J. 1987. *Occupational social work*. Englewood Cliffs, NJ: Prentice Hall.

Groenveld, I. F., Proper, K. I., Absalah, S., van der Beek, A. J. & van Mechelen, W. 2011. An individually based lifestyle intervention for workers at risk for cardiovascular disease: A process evaluation. *American Journal of Health Promotion*, 25:396–401.

Grover, F., Wu, H.D., Blanford, C., Holcomb, S. & Tidler, D. 2002. Computer-using patients want internet services from family physicians. *Journal of Family Practice*, 51:570–572

Guest, G., MacQueen, K.M. & Namey, E.E. 2019. *Introduction to Applied Thematic Analysis.* Thousand Oaks: SAGE Publications.

Haber, Y., Levkovich, I., Hadar-Shoval, D. & Elyoseph, Z. 2024. The artificial third: A broad view of the effects of introducing generative artificial intelligence on psychotherapy (Preprint). *JMIR Mental Health*. 11:e54781.

Hanley, T. & Reynolds, D.J. 2009. Counselling Psychology and the internet: A review of the quantitative research into online outcomes and alliances within text-based therapy. *Counselling Psychology Review*, 24(2):4-13.

Harari, Y.N. 2019. *21 Lessons for the 21st Century*. London: Vintage.

Harari, Y.N. 2016. *Homo Deus: A Brief History of Tomorrow*. London: Vintage.

Hasson, F. & Keeney, S. 2011. Enhancing rigour in the Delphi technique research. *Technological Forecasting & Social Change,* 78:1695-1704.

Hasson, F., Keeney, S. & McKenna, H. 2000. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4):1008-1015.

Hoffer, J. 2002. *Modern system analysis and design*. 4th ed. New Jersey: Pearson Prentice.

Howick, J., Morley, J. & Floridi, L. 2021. An Empathy Imitation Game: Empathy Turing Test for Care- and Chat-bots. *Minds and Machines,* 31:457-461, art. #21. doi:10.1007/s11023-021-09555-w

Hsu, J.Y.J & Hasmath, R. 2017. A maturing civil society in China? The role of knowledge and professionalisation in the development of N.G.O.s. *China Information*, 31(1):22-42, 2017.

Hsu, C. & Sandford, B. 2007. The Delphi Technique: Making Sense of Consensus. Practical Assessment, Research & Evaluation, 12, 1-8.http://pareonline.net/pdf/v12n10.pdf

Jafari, H., Kassan, A., Reay, G., & Climie, E. A. 2022. Resilience in Refugee Children and Youth: A Critical Literature Review. *Canadian Psychology/Psychologie canadienne,* 3, art. #09. doi: 10.31703/glr.2018(III-I).09

Jirava, P. 2004. System development life cycle. *Industrial Management & Data Systems,* 109(3): 338-356.

Johnson, R.B. & Gray, R. 2010. A history of philosophical and theoretical issues for mixed methods research. In: Tashakkori, A. & Teddlie, C., eds. *Handbook of mixed methods in social and behavioural research*. Thousand Oaks, CA: Sage. pp. 69-94.

Kaniki, A.M. 2006. Doing an information search. In: Terre Blanche, M., Durrheim, K. & painter, D., eds. *Research in practice*. Cape Town: Cape Town University Press. pp. 18-33.

Kasi, V., Keil, M., Mathiassen, L., & Pedersen, K. 2008. The post mortem paradox: A Delphi study of IT specialist perceptions. *European Journal of Information Systems*, 17:62-78.

Kenny, M.C., & Mceachern, A. G. 2004. Telephone counseling: Are offices becoming obsolete? *Journal of Counseling & Development*, 82:199–203.

Kettner, P.M., Moroney, R.M. & Lawrence, L.M. 2008. *Designing and Managing Programs.* 3rd ed. Thousand Oaks: Sage.

Krishnav, D. 2022. *Basics of building an Artificial Intelligence Chatbot* [Blog post]. https://www.mygreatlearning.com/blog/basics-of-building-an-artificial-intelligence-chatbot/#textchatbot Date of access: 01 May. 2022.

Kirsten, N. 2022. Big data application in AI development. [Telephonic interview]. 2 May, Krugersdorp.

Kruger, A. & Van Breda, A.D. 2001. Military social work in the South African National Defence Force. *Military Medicine,* 166(11):947-951.

Levinson, C. 2018. *The Definition of Lean Operations.* https://bizfluent.com/about-6666969-definition-lean-operations.html Date of access: 19 Mar. 2022.

Lexalytics. 2020. *The 7 Basic Functions of Text Analytics & Text Mining.* *https://www.lexalytics.com/lexablog/text-analytics-functions-explained* Date of access: 30 Apr. 2022.

Li, S., Long, F., Liu, S., Lee, Q., Guo, R., Li, C. & Liu, Y. 2006. Herbs for medicinal baths among the traditional Yao communities of China. *Journal of ethnopharmacology*, 108:59-67.

Linstone, A.H. & Turoff, M. 1975. *The Delphi Method: Techniques and Applications*. Boston: Addison-Wesley.

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. 2013. I can help you change! An empathic virtual agent delivers behavior change health interventions. *A.C.M. Transactions on Management Information Systems (TMIS)*, 4(4):19.

Luther, A. 2017. The Entity Mapper A: Data Visualization Tool for Qualitative Research Methods. *Leonardo*, 50(3):268-271.

Mallen, M.J. & Vogel, L.V. 2005. Online Counseling: A Need for Discovery. *The Counseling Psychologist,* 33(6):910-921.

McGee, V. 2022. *Virtual Therapy vs. In-Person Therapy.* https://www.psychology.org/resources/virtual-therapy-vs-in-person/ Date of access: 06 Apr. 2022.

Mead, D. & Mosely, L. 2001. The use of the Delphi as a research approach. *Nurse researcher*, 8(4):4-23.

Moore, M.M. & Ophir, Y. 2021. Big Data, Actually: Examining Systematic Messaging in 188 Romantic Comedies Using Unsupervised Machine Learning. *Psychology of Popular Media*, 14(3), art. #349. doi:10.1037/ppm0000349

Mostert-Phipps, N. Pottas,D. & Korpela, M. 2013. A South African perspective on factors that impact the adoption and meaningful use of health information technologies. *S Afr Fam Pract,* 55(6):545-554.

Munn, Z., Peters, M.D.J., Stern, C., Tufanaru, C., McArthur, A. & Aromataris, E. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *B.M.C. Med Res Methodol*, 18:143.

Neuman, W. 2003. Social Research Methods: Qualitative and Quantitative Approaches. Essex: Pearson.

Nagarajan, M. & Yuvaraj, S. 2021. Mental health counsellors' perceptions of the use of technology in counselling. *Current Psychology*, 40:1760-1766, art. #4. doi:10.1007/s12144-018-0104-4

Novotney, A. 2017. A growing wave of online therapy. *Monitor on Psychology*, 48(2):48.

Ogbeifun, E., Agwa-Ejon, J., Mbohwa, C. & Pretorius, J.H.C. 2016. The Delphi technique: A credible research methodology. *Industrial Engineering and Operations Management,* 8-10.

Patterson, D.A. & Cloud, R.N. 1999. The application of artificial neural networks for outcome prediction in a cohort of severely mentally ill outpatients. *Journal of Technology for Human Services*, 16(3):47-61.

Paul, J. & Criado, A.R. 2020. The art of writing literature review: What do we know and what do we need to know? *International Business Review*, 29(4):101-108.

Pinciroli, F., Justo, J.L.B. & Forradellas, R. 2022. Systematic mapping study: On the coverage of aspect-oriented methodologies for the early phases of the software development life cycle. *Journal of King Saud University-Computer and Information Sciences*, 34(6):2883-2896.

Poh Li, L., Rafidah, A.M.J. & Abdullah, H.S. 2013. Understanding the two sides of online counseling and their ethical and legal ramifications. *Procedia - Social and Behavioral Sciences*, 103:1243-1251.

Potts, C., Ennis, E., Bond, R.B., Mulvenna, M.D., McTear, M.F., … O'Neill, S. 2021. Chatbots to support mental wellbeing of people living in rural areas: Can user groups contribute to co-design? *Journal of Technology in Behavioral Science*. 6(4):652–665.

Ray, P. & Chakrabarti, A. 2017. *Twitter sentiment analysis for product review using lexicon method.* Paper delivered at the International Conference on Data Management, Analytics, and Innovation (ICDMAI), India, Pune. https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/apirference/get-search-tweets Date of access: 30 Apr. 2022.

Roestenburg, W.J.H. 2021. *Analysing technology acceptance by social workers in South Africa*. Potchefstroom: North-West University. (Dissertation – M.B.A.).

Rothman, J. & Thomas, E.J. 1994. *Intervention Research: Design and Development for Human Service*. New York: Haworth Press.

Rowe, G., Wright, G. & McColl, A. 2005. Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence. *Technology Forecast and Social Change*, 72:377- 399.

Rubin, A. & Babbie, E. 2016. *Essential Research Methods for Social Work*. 4th ed. Boston, MA: Cengage Learning.

Rubin, A. & Babbie, E. 2013. *Essential Research Methods for Social Work*. 3rd ed. Boston, MA: Cengage Learning.

Rudestam, K.E. & Newton, R.R. 1992. *Surviving your dissertation: A comprehensive guide to content and process*. California: Sage.

Saliba, Y. & Barden, S. 2017. Counselors and Workplace Wellness Programs: A Conceptual Model. *The Professional Counselor*, 7(2):104-113.

Schurink, W.J., Schurink, E.M. & Fouché, C.B. 2021. Qualitative data analysis and interpretation. In: Fouche', C.B., Stryfom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 391-415.

Schlenger, W., Jöllenbeck, M., Stamer, T., Grosse, A. & Ochsmann, E. 2022. Digitising Social Counseling—Insights for Workplace Health Management. *International Journal of Environmental Research and Public Health*, 19:1-16.

Schwab, K. 2016. *The Fourth Industrial Revolution*. Geneva: World Economic Forum.

Schwab, K. 2015. *The Fourth Industrial Revolution: what it means, how to respond.* *https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/* Date of access: 11 Mar. 2022.

Schuetzlera, R.M., Mark, G. & Giboneya, J.S. 2020. Impact of Chatbot Conversational Skill on Engagement and Perceived Humanness. *Journal of management information systems,* 37(3): 875-900.

Shah, S.R., Ahmed, F. & Khan, R. 2018. Writing a Critical Review of Literature: A Practical Guide for English Graduate Students. *Global Language Review*, 3:136-153.

Skinner, R., Nelson, R.R., Chin, W.W. & Land, L. 2015. The Delphi Method Research Strategy in Studies of Information Systems. *Communications of the Association for Information Systems,* 37:31-63.

Skjuve, M.B. & Brandtzaeg, P.B. 2018. Chatbots as a new user interface for providing health information to young people. https://sintef.brage.unit.no/sintef-xmlui/ Date of access: 05 Oct. 2021.

Sperandeo, R., Cioffi, V., Mosca, L.L., Longobardi, T., Moretto, E., Alfano, Y.M., … Maldonato, N.M. 2021. Exploring the Question: Does Empathy Work in the Same Way in Online and In-Person Therapeutic Settings? *Front Psychol.,* 21(12), art. #671790. doi: 10.3389/fpsyg.2021.671790

Strydom, H. 2021. Sampling techniques and pilot studies in qualitative research. In: Fouché, C.B., Strydom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 379-390.

Sullivan, G.M. 2011. A primer on the validity of assessment instruments. *J Grad Med Educ*. 3:119-20.

Sunter, C. & Ilbury, M. 2021. *Thinking the future: New Perspectives from the shoulders of giants.* Cape Town: Penguin Books.

Susanti, A.R., Djanta, T.& Kusuma, W.A. 2017. Twitter's Sentiment Analysis on Gsm Services using Multinomial Naïve Bayes. *Telkomnika,* 15(3):1354-1361.

Tambe, M. & Rice, E. 2018. Merging Social Work Science and Computer Science for Social Good. In: Tambe, M & Rice, E., eds. *Artificial Intelligence and Social Work*. Cambridge: Sheridan Books. pp. 3-16.

Terrell, S. R. 2012. Mixed-Methods Research Methodologies. *The Qualitative Report*, *17*(1):254-280.

Tongco, C.D. 2007. Purposive sampling as a tool for informant selection. *Ethnobotany research & applications,* 5:147-158.

Umoquit, M., Tso, P., Varga-Atkins, T., O'Brien, M., & Wheeldon, J. 2013. Diagrammatic elicitation: Defining the use of diagrams in data collection. *The Qualitative Report*, 18(30):1-7.

Van Ewijk, H. 2018. *Complexity and social work.* Abingdon: Routledge.

Van Ewijk, H. 2010. *European social policy and social work. Citizen-based social work*. Abingdon: Routledge.

Venkatesh, V., Brown, S.A. & Sullivan, Y.W. 2016. Guidelines for Conducting Mixed-methods Research: An Extension and Illustration. *Journal of the Association for Information Systems*, 17(7):436-495.

Wahyono, I.D., Fadlika, I., Asfani, K., Putranto, H. & Hammad, J. 2019. New adaptive intelligence method for personalised adaptive laboratories. *Electrical, Electronics and Information Engineering,* 6:196-200.

Walliman, N. 2011. *Research Methods the basics*. London: Routledge.

Walter, D. & Ophir, Y. 2019. News Frame Analysis: An Inductive Mixed-method Computational Approach. *Communication Methods and Measures*, 13(4):248-266.

Xu, M & David, J.M. & Kim, S.H. 2018. The Fourth Industrial Revolution: Opportunities and Challenges. *International Journal of Financial Research,* 9(2):90-95.

Xu, S., Li, Y. & Zheng, W. 2017. Bayesian Multinomial Naïve Bayes Classifier to Text Classification. *Lecture Notes in Electrical Engineering,* 448:347-352.

Zhang, Q. 2022. A Big Data-Driven Approach to Analyse the Influencing Factors of Enterprise's Technological Innovation. *Computational Intelligence and Neuroscience*, 2022, art. #3785685. doi:10.1155/2022/3785685

CHAPTER 2: LITERATURE STUDY

2.1. INTRODUCTION

> *"Making it up?" said Marvin, swivelling his head in a parody of astonishment, "Why should I want to make anything up? Life's bad enough as it is without wanting to invent any more of it."*
>
> (Adams, 1979:92).

Marvin the Paranoid Android is a fictitious character that Douglas Adams created in the *Hitchhiker's Guide to the Galaxy* novels. Marvin is the android on board the spaceship, Heart of Gold. He was developed as one of the Genuine People Personalities (GPP) technical prototypes that Sirius Cybernetics Corporation ultimately abandoned. He has a "brain the size of a planet" that he rarely, if ever, gets to use, which contributes to his tremendous unhappiness and boredom (Adams, 1979:81). He is only required to execute simple tasks, such as opening doors for his team. Since no task that could be given to Marvin could consume even a small percentage of his great intelligence, his life is profoundly dismaying to him, as he claims to be 50,000 times more intelligent than a human. Although fictional (if not a little foolish), Marvin (Adams, 1979; Adams, 1980; Adams, 1982) provides a glimpse into how academics and futurists envisaged the future use of AI and ML. The interest in *thinking machines*, even in popular culture, received great attention after Turing's machines successfully deciphered encrypted German communication during the Second World War.

This interest led to a small group of scientists convening at Dartmouth College in 1956. This group aimed to investigate methods to teach machines to learn. The theory and mathematics investigated during the Dartmouth Workshop supported the notion that machines could indeed learn – thus a new field of study and application was born (Chivers, 2019; McCarthy *et al.*, 1956).

Marvin does not sound so far-fetched in 2023 with the emergence of ChatGPT (Chat Generative Pre-Trained Transformer), which was launched in June 2022 by OpenAI Inc., an American AI research laboratory. With investments from multiple tech firms, it cost an estimated 50 billion United States Dollars (USD) to develop the chatbot, which is currently running for USD100,000 per month. ChatGPT is currently opening up never-before-seen avenues for AI and is undoubtedly causing a media frenzy since it managed to write a 10,000 word article in 30 seconds, pass the Minnesota University Law School exam, produce academic journal quality papers and pass various other exams (AFP Wire Service, 2023; Lucey & Dowling, 2023; Nolan, 2023; OpenAI, 2023). The accuracy and intelligence of ChatGPT have necessitated various universities and schools to ban its use (Johnson, 2023; Lipman & Distler, 2023).

Chatbots are computer programs that engage with users in their natural language via text or voice to provide access to services and information. Several entities are developing chatbots to automate customer service and offer universal access to services. Conversational agents (software programs that communicate with people using natural language) have recently attracted increasing interest. Reference has been made to a "chatbot tsunami", as text-based chatbots, which allow human-machine contact via natural written language, have become more prevalent during the past 10 years (Grudin & Jacques, 2019:1).

One of the most enduring objectives of AI is the creation of social chatbots - or intelligent dialogue systems - that can have sympathetic dialogues with people. Early conversational systems such as Eliza (Weizenbaum, 1966), Parry (Colby *et al.*, 1971), and Alice (Wallace, 2009), were created to replicate human behaviour in a text-based conversation and, as a result, pass the Turing Test (Turing, 1950) within a defined context. Eliza, Parry and Alice were chatbots that were designed within and for use within the domain of psychology. No entities such as the mentioned agents exist within the domain of social work. Tambe and Rice (2018), therefore, challenge social workers to make a radical shift in their social work perspective by imagining how AI can benefit client populations.

The hypothesis is that the development of an AI-generative chatbot counselling framework within the field of social work and EHWPs will lead to a modernisation of the social work knowledge base, the discipline, and the services provided by social workers, and promote the future relevance of the profession in a technologically enhanced world. The aim is to provide insights that can inform the future development of mental health services and AI applications in social work by exploring the key factors that contribute to the successful development and implementation of such a framework.

The goal of this study is to create and evaluate a multiple-response AI algorithm to support a conceptual AI generative chatbot counselling framework as part of a larger social work EHWP. This literature study aims to examine the following elements to achieve the overarching goal:

- AI chatbots:
  - Benefits of AI chatbots and virtual therapeutic services;
  - Limitations of AI chatbots and virtual therapeutic services; and
  - Holistic features of AI chatbots.
- Methodologies deployed in the development of AI chatbots;
- The use of big data in the development of AI chatbots;
- The use of linguistics in the development of AI chatbots;
- The use of AI chatbots within social work and the EHWP context;
- Clinical theoretical framework; and

- Ethical considerations in the deployment of AI in social work and an EHWP.

2.2. LITERATURE STUDY METHODOLOGY

A literature study places the research topic in context by describing how it relates to a particular field. A literature study aims to uncover knowledge gaps, assist in the formulation of research problems, identify theoretical frameworks, problems, and variables pertinent to the study topic, as well as identify conceptual and operational definitions. The process outlined by Jafari *et al.* (2022) was adjusted to produce a sound literature study. This procedure involved i) locating information in databases using search phrases; ii) excluding any materials that did not subscribe to the chosen timeframes; iii) omitting content devoted to other topics; iv) manually locating pertinent materials in the reference lists of the materials already included; v) reviewing the final information included for review; vi) examining texts, and finally vii) reporting on the findings.

2.3. RATIONALE FOR THE STUDY

The IoT, which connects physical and digital applications, is one of the driving forces behind the notion of the 4IR. At its most basic level, IoT can be defined as a relationship between commodities and people that is made possible by interconnected technology and a variety of platforms (Giglio, 2021; Hajjaji *et al.,* 2021). Ashton (2009) establishes the essential concepts of IoT by introducing device-to-device connectivity on a larger scale than had previously been possible. In terms of semantics, Ashton (2009:97) describes the IoT as a "global network of interconnected objects". Patel *et al.* (2016) note that the IoT is a dynamic global network infrastructure that can identify, control, and monitor systems via the Internet according to a specific agreement protocol and through the interconnection of physical and virtual things based on ICT interoperability. Examples of the IoT in our daily lives include wearable devices, self-driving cars, crop monitors, logistic tracking and retail inventory tracking (Shreyas *et al.,* 2020).

Customers, people or businesses are at the centre of the IoT, which strives to enhance how these groups provide and receive services. To maximise the value of their assets, organisations are managing and appreciating them differently due to technological advancements. According to Javaid *et al.* (2022), the combination of AI and IoT has the potential to produce powerful new applications that can improve daily life in a variety of ways. As AI and IoT continue to advance, one should expect to witness even more inventive and ground-breaking applications in the coming years. The adoption of AI is required by new operating models, which necessitate businesses and organisations to reassess how they now operate to function more agilely and effectively. As with other scientific fields, social work and EHWPs have been impacted by technological advancements. Fortunately, technology supports mental health in various ways.

Although not yet locally applied, the incorporation of big data technologies into social work is actively being made possible abroad in fields such as patterns in homelessness, healthcare and social services planning, predictive modelling in risk assessment, as well as program evaluation and evidence-based practice (Foster *et al.,* 2020). Big data-driven social work will have an application in the future that will reflect how it can capture, update, and respond to real-time data on human behaviour and needs (Fava *et al.*, 2016; Marsch, 2021).

AI and social work may seem to be unrelated fields; however, Tambe and Rice (2018) contend that by combining these two fields, a dynamic new science can be created. According to the authors, the collaborations that are exclusive to both fields make AI4SG or AI and social work, a new intellectual arena. Given current developments in AI and DL, chatbots are a promising supplement to conventional forms of service.

This study suggested adding chatbot contact to traditional practice rather than proposing that AI technology and mental healthcare professionals be replaced by it. The field of social work is, however, lacking in terms of the 4IR, automation, AI, and the IoT because there are no AI social work-specific AI therapy platforms available. It makes sense that social work adapts or at the very least expands, its problem-solving methodology in the face of the 4IR, automation, AI, and the IoT, with digital-led technologies rapidly approaching a tipping point of becoming more widely available and cost-effective (Sunter & Ilbury, 2021).

Chatbots may be the solution since they offer a novel, cutting-edge, interactive modality of connecting with individuals who are using mental health services. The knowledge base and theory of social work would be enriched by this study. If the study were successful, it could assist in modernising the field of social work, which may enable it to stay relevant to the 4IR, automation, and the IoT. The long-term goal of this study was, therefore, to enable users of mental health EHWP services to obtain services via a chatbot based on the framework created in this study.

2.4. ARTIFICIAL INTELLIGENCE (AI) CHATBOTS

AI is a broad concept, with various interpretations depending on the setting, the period, and the intended use (Visvikis *et al.*, 2019). It is believed that this academic field was established in 1956 by computer scientist John McCarthy (McCarthy, 1956) at the Dartmouth Workshop (Moor, 2006). Moor (2006) expanded on McCarthy's conceptualisation of AI and described it as intelligent behaviour shown by machines, as opposed to the innate intelligence characteristic of humans and animals, as a relatively broad description.

AI systems can be analytical, human-inspired or humanised (Kaplan & Haenlein, 2019). Only cognitive intelligence-related traits are possessed by analytical systems, which use learning based on prior experience to create predictions. AI systems with human inspiration have cognitive components as well as emotional intelligence and understanding. AI systems that have been

"humanised" can display social, emotional, and cognitive intelligence, as well as self-consciousness and self-awareness in interactions with humans. The development of social networks and other platforms, cloud storage, computing, and other factors have made it easier for people to access ever-larger amounts of data for learning, which has benefited AI techniques in the 21st century. According to Molano *et al.* (2018), these factors also include the availability of the algorithms and libraries. Older concepts and theories can now be applied to solve issues and complete jobs in the real world, even by non-experts using systems that have already arrived in the industry. Interacting with people, addressing their concerns, and resolving their challenges are some of the primary functions of online chatbots. Since these agents offer more interactive message delivery to users, specifically addressing their questions, the experience they provide is thought to be superior to static information delivery, such as a list of frequently asked questions, (Go & Sundar, 2019). Several companies have started using chatbots in place of human agents or telephone-based call support.

AI and social work may appear to be distinct disciplines; however, Tambe and Rice (2018) propose that integrating these fields has the potential to establish a transformative and impactful area of scientific inquiry. Given recent developments in AI and DL, the ubiquitous use of online digital devices, and the advent of enormous data sets from digital trails, chatbots are a viable supplement to traditional varieties of social work services. With the development and examination of intelligent hardware and software, also known as intelligent agents, AI is gradually integrated into our daily lives. Intelligent agents can perform a wide range of jobs, from manual labour to complex procedures (Mindell & Reynolds, 2023). A chatbot is one of the simplest and most common forms of intelligent human-computer interaction (HCI) and a classic example of an AI system. In its most simple form, a chatbot is a computer program that responds intelligently when spoken to or written to and uses NLP to comprehend and respond to human interactions (Bansal & Khan, 2018; Khanna *et al.*, 2015).

As with every new technological solution or methodology, some industries may adopt chatbots before others. A significant portion of the first use or engagement may be a result of factors such as competitiveness, cost savings or strategic advantage. Several sectors such as healthcare, business, hospitality, banking/financial, retail service, and publishing, stand to gain significantly from the usage of chatbots, which has made the case for early investors in this technology (Mierzwa *et al.*, 2019). In several circumstances, social chatbots are being regarded to automate verbal social interaction and minimise human presence. Chatbots have been tested or evaluated through surveys or deployment in a variety of settings with differing levels of readiness (Heath *et al.*, 2019).

This literature study aimed to examine available resources regarding AI chatbots. The research sought information that had specific reference to AI chatbots' potential benefits, potential

limitations, holistic features, the use of big data and linguistics in its development, as well as the use of AI chatbots within the social work and EHWP context.

2.4.1. Benefits of Artificial Intelligence (AI) chatbots and virtual therapeutic services

Before examining the potential benefits of AI chatbot therapeutic services, it seems fitting to examine the benefits of online or virtual therapeutic services, which are delivered by humans via platforms such as live chat. An interesting study by Fang *et al.* (2018) concludes that the use of online counselling services, whether in conjunction with in-person counselling or exclusively through online sessions, was reported by their study participants as having several benefits. Accessibility and a sense of anonymity were two of the main benefits identified. Several participants chose the text-based online counselling platform instead of face-to-face counselling because they found it difficult to schedule an in-person meeting to receive face-to-face services. The participants furthermore appreciated the flexibility of an online service for users who were occupied with studies, jobs, and personal responsibilities or dealing with personal struggles or mental health issues that would have otherwise made it challenging to attend counselling. Participants expressed appreciation for the flexibility offered by online counselling, especially regarding scheduling or transportation issues. Many clients voiced support for the ability to immediately record their emotions rather than having to wait until the next scheduled counselling appointment (Fang *et al.*, 2018).

Zumstein and Hundertmark's (2017) study outlines the advantages of AI chatbots for both the provider and the user. For the provider, the research outcomes outline the potential benefits as 24-hour-per-day services that can be accessed anywhere and at any time; the establishment of new and direct user contact points; the facilitation of new methods and types of data collection; the collection and storage of high volumes of user data; personalisation and automation of communication of services; reduced service and support costs; high convenience and ease of use, as well as the potential to save time and costs.

Adamopoulou and Moussiades (2020) state that users mostly perceive chatbots as friendly companions and not as mere assistants. Ahmad *et al.* (2018) note that if a chatbot is sufficiently established and deployed into use, it may function as a tool to encourage user interaction and offer a positive user experience between users and the specific domain being served. Chatbots have the potential to be helpful tools for people with mental disorders, particularly for those who are hesitant to seek mental health assistance out of fear of stigmatisation. There have been various studies exploring the use of chatbots in the context of mental health, such as publications by Balcombe and De Leo (2022), Boucher *et al.* (2021), Denecke *et al.* (2021) and Bridgman and Lazar (2019). However, to inform mental health professionals and potential users about the key features of chatbots and their potential applications, as well as to identify major gaps in the

existing literature and guide future research, there is a need to systematically compile and analyse the available evidence (Abd-alrazaq *et al.*, 2019).

Wu *et al.* (2020) have identified additional benefits of AI chatbots as the ability to render services to multiple users simultaneously. The expense of human labour and the strain of providing human services might be significantly reduced by using a chatbot, which could in turn improve the customer experience and improve near-instant response times.

Numerous benefits of online mental health services have been identified when considering the counselling or therapeutic application of chatbots and virtual counselling, as deployed in psychology. These benefits include the elimination of the stigma that has historically been associated with mental healthcare, ease of access and convenience, privacy, comfort and safety, ease of communication, cost- and general effectiveness, improved consistency, more frequent sessions, better progress, flexible sessions, increased access, anonymity, enhanced autonomy, the immediacy of services and, finally, the motivation to attend in-person services, should the user choose to do so (McGee, 2022).

2.4.2. Limitations of Artificial Intelligence (AI) chatbots and virtual therapeutic services

Fang *et al.* (2018) have studied the distinctive limitations or challenges of online or virtual therapeutic services, which were delivered by humans via digital platforms. They state that the most frequent criticism was that many participants experienced difficulties using the secure online counselling software due to technical and procedural issues. Participants reported trouble navigating the online experience and managing administrative concerns, among other technological obstacles. Online sessions have occasionally been terminated as a result of connection issues. Several participants emphasised how difficult it was to remember their passwords, figure out new software, and log in. Some participants indicated an overall reluctance, annoyance or unease with text-based online counselling, including the online format and procedure.

Zumstein and Hundermark (2017) make note of certain weaknesses or risks in their study, which are categorised as possible disadvantages for the provider as well as for the user. For providers of AI chatbots, possible disadvantages have been identified. These include malfunctioning chatbots, unanswered questions or irrelevant responses, high costs of investment in IT infrastructure and tools, extension of IT and analytics architectures, lack of awareness and acceptance of users and potential users, information security or cyber security risks, provider image and reputational risks.

Zumstein and Hundermark (2017) have identified risks for users of AI chatbots as protection of data and sensitive personal information concerns, lack of understanding and experience when conversing with an AI chatbot, biased personalised information, social isolation, and ethical

concerns, as well as new communication styles that may require some time for users to adjust to. The authors also note that in the transition phase, the majority of traditional (offline and online) channels still need to be offered, and users need to be encouraged to use new tools and technology.

Wu *et al.* (2020) mention further shortcomings of AI chatbots as learning issues that may arise from the fact that the datasets and knowledge bases now available are still in their infancy, which is a major restriction for any form of chatbot. According to Kumar *et al.* (2016), the absence of important medical concepts and the presence of irrelevant medical concepts can also result in information loss and low precision in local data mining methodologies. Santos *et al.* (2022) focus on the learning style of a specific AI chatbot and note that if certain styles of learning are employed, the chatbot could learn out-of-domain knowledge and provide incorrect feedback while regarding the feedback as being correct.

In terms of user requirements and experience, Valério *et al.* (2017) state that a chatbot manages many diverse user types. New users are unaware of the capabilities of a specific chatbot that they are dealing with. This may lead to an explorer user who will test sentences to determine if the chatbot can understand them and will leave if the chatbot displays restricted responses and content. Some users are unwilling to experiment with technology out of fear or reluctance or they may be averse to new methods of carrying out certain jobs.

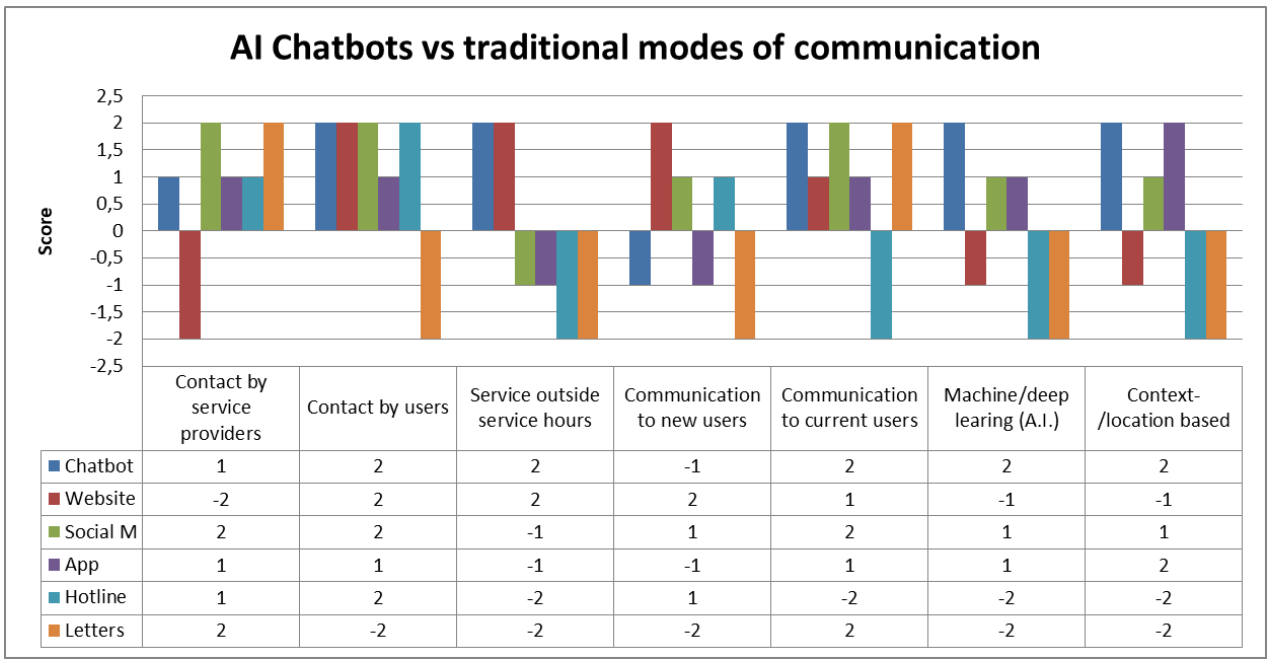
Conversations between humans and chatbots differ significantly from human-to-human conversations in terms of both content and quality (Adamopoulou & Moussiades, 2020). In an evaluation of the design of chatbots, Jain *et al.* (2018) conclude that service users believed that chatbots frequently failed to understand their input text or discern their intentions, making it difficult for chatbots to interact with or respond to them effectively. The participants' first-time interactions with eight pre-selected chatbots over three days and nearly 10,000 messages demonstrated that their expectations were not met due to the chatbots' poor natural language abilities and their sparse feature set. Participants indicated that they were either disappointed or frustrated. Users were unable to judge the AI of the bots; this is consistent with Luger and Sellen's (2016) findings. The participants did not anticipate using the majority of the chatbots in the future due to the limited set of functionalities supplied by the chatbots. Jones *et al.* (2022) add that users want to feel as though they are getting a real experience; therefore, if users do not feel that an online discussion is genuine, it can undermine the advantages of adopting new technologies.

As an additional note of concern, Mierzwa et al. (2019) mention that healthcare industries must ensure that automation does not lead to any unexpected problems when new technology is introduced. The rise of harmful bots poses a significant threat to individuals and industries, with the potential for social, psychological, economic, democratic, and legal harm (Daniel *et al.*, 2019). Among industries that use automation, the healthcare sector is ranked fourth in susceptibility to

malicious bot attacks. Rivas *et al.* (2018) state that users will hold the person who created the chatbot responsible if the chatbot engages in ethically repugnant behaviour, while the chatbot will also bear some of the responsibility but on a smaller scale of culpability. Users seem to hold chatbots to some of the same moral standards as humans. AI chatbots offer unique opportunities to improve service delivery in social work and EHWPs. To ensure fair and responsible implementation, the integration must be guided by robust ethical considerations. By focusing on fairness, data protection, and ethical AI development, the South Africa National Artificial Intelligence Policy Framework serves as a foundational step towards developing a comprehensive national policy. These preliminary guidelines are essential for resolving bias concerns and ensuring that AI applications in social work are in line with international AI governance standards as well as national socioeconomic priorities (Department of Communications and Digital Technologies, 2024).

2.4.3. Holistic features of Artificial Intelligence (AI) chatbots and virtual mental health therapeutic services

Whether embodied or not, therapeutic chatbots are not inert objects. They are interactive computing systems created to alter people's opinions or behaviours, so they should be regarded as persuasive technologies. According to research by Hiland (2021), humans can have feelings towards technologies that resemble their feelings for and treatment of other people. A study undertaken by Zumstein and Hundertmark (2017) compares AI chatbots to traditional modes of mental health service delivery. The authors tabled their findings regarding the efficiency of chatbots versus other modes of service delivery, which the researcher adapted and illustrated in Figure 10.



	Contact by service providers	Contact by users	Service outside service hours	Communication to new users	Communication to current users	Machine/deep learing (A.I.)	Context-/location based
Chatbot	1	2	2	-1	2	2	2
Website	-2	2	2	2	1	-1	-1
Social M	2	2	-1	1	2	1	1
App	1	1	-1	-1	1	1	2
Hotline	1	2	-2	1	-2	-2	-2
Letters	2	-2	-2	-2	2	-2	-2

Figure 10. AI chatbots versus traditional modes of communication (adapted from Zumstein and Hundertmark, 2017)

Zumstein and Hundertmark (2017) compare service delivery via AI chatbots to traditional modes of service delivery, with specific reference to websites, social media, applications, hotlines, and letters in Figure 10. From this figure, it can be deduced that AI chatbots possess various strengths concerning service delivery but lack in the domains of initiating contact with service users and communication with new or potential service users.

A scoping review by Abd-alrazaq *et al.* (2019) provides an overview of the key features of mental health-specific AI chatbots (Table 4). The authors have identified 53 studies (*n*=53) that assessed 41 different mental health chatbots. The most common use of the studied chatbots was for the delivery of psychological therapy, training, and screening. Of these, 10 chatbots were based on CBT, and eight were specifically designed for people living with depression and anxiety. The findings conclude that the majority of chatbots are rule-based, implemented in standalone software, start the conversation, contain embodiment, target anxiety and depression, and are utilised in developed countries (see Table 4 for a summary of AI chatbot features). The lack of mental health chatbots in less developed countries presents a unique opportunity for the field of social work to broaden its footprint by bridging existing gaps with technology.

Furthermore, in 70% of studies assessed by Abd-alrazaq *et al.* (2019), chatbots were deployed on standalone software, while the remaining 30% were used on web-based platforms. The bulk of research (92.5%) used decision trees or predetermined rules to guide chatbots' response generation (rule-based). The remaining four experiments used chatbots that produced their responses using ML techniques. Supervised ML was utilised in one study, whereas reinforcement

learning was used in the other. In 86.8% of the trials, chatbots took the initiative to start the conversation, while in the remaining seven research studies, both chatbots and users could. Written language via keyboard and mouse (*n*=26), spoken language via microphone (*n*=8), a combination of spoken and visual languages via microphones and camera or Kinect (*n*=10), a combination of written and spoken languages (*n*=7), and a combination of written and visual languages were all possible ways for users to communicate with the chatbots. In contrast, chatbots interacted with users using modalities of a combination of written text, spoken, and visual languages (*n*=28), a combination of spoken and visual languages (*n*=11), just written language (*n*=9), and a combination of written and visual languages (*n*=5). Virtual representations, such as avatars or virtual humans, were embodied in chatbots in 44 investigations (83%). The chatbots under investigation were deployed to provide assessments and services for a range of psychological disorders that include depression, autism, post-traumatic stress disorder (PTSD), uncategorised mental health disorders, anxiety, substance abuse disorders, schizophrenia, dementia, phobia, stress and eating disorders.

Table 4. Features of mental health AI chatbots (adapted from Abd-alrazaq *et al.*, 2019)

Characteristic	AI chatbot	No. of studies
Purpose	Therapy	17
	Training	12
	Screening	10
	Self-management	7
	Counselling	5
	Education	4
	Diagnosing	2
Platform	Stand-alone software	37
	Web-based software	16
Response generation	Rule-based	49
	Artificial intelligence [AI]	4
Dialogue initiative	Chatbot	46
	Both	7
	User	0
Input modality	Written	26
	Spoken	8
	Spoken & visual	10
	Written & spoken	7
	Written & visual	2
Output modality	Written	9
	Spoken	0
	Spoken & visual	11
	Written, spoken & visual	28

	Written & visual	5
Targeted disorders	Depression	16
	Autism	10
	PTSD	7
	Mental disorders	7
	Anxiety	7
	Substance use disorders	5
	Schizophrenia	3
	Dementia	3
	Phobia	2
	Stress	2
	Eating disorders	1

2.5. METHODOLOGIES DEPLOYED IN THE DEVELOPMENT OF AI CHATBOTS

Numerous methodologies can be deployed in the development of AI chatbots and frameworks. The most commonly utilised methodologies are discussed in this section.

2.5.1. Characteristics for consideration in the design of an Artificial Intelligence (AI) chatbot framework

Bendig *et al.* (2019) note that existing psychology chatbots have been developed based on characteristics that typically include i) an area of use (prevention, treatment, follow-up); ii) underlying approaches (CBT, psychodynamic approaches); iii) technical implementation (decision tree, ML); iv) goals/endpoints (improvement of symptoms, support for behavioural change), and v) performance (text-base, speech-based, rational/contextual). To inform the chatbot framework, which will then be used to construct a computational model for interaction, a corpus or dataset for analysis, is required. The previously mentioned information is utilised in order to inform the architecture of the chatbot, which consists of textual responses, NLP, response generation, knowledge base creation, and dialogue management (DM) (Cahn, 2017). This commonly used process is displayed in Figure 11.

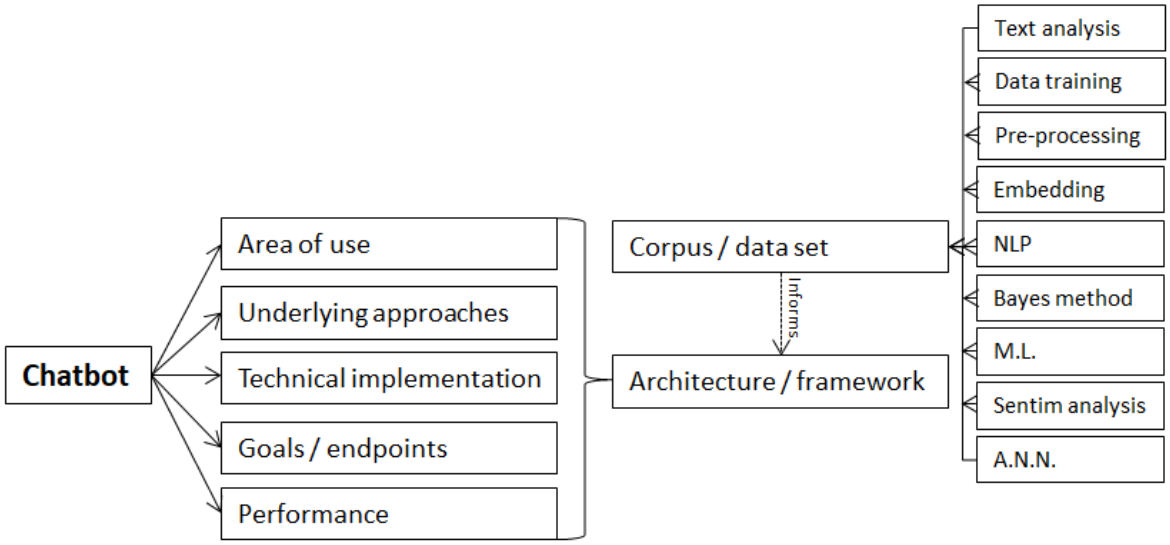


Figure 11. Commonly deployed strategies in chatbot framework development

2.5.2. Text analysis

NLP, a method that provides computers with the ability to comprehend human speech or text and relate to it, is frequently combined with ML, a technology that deals with providing computers with the ability to learn. Specifically, NLP has broadened its study scope to encompass the entirety of phrases rather than just a single word. This makes it feasible to handle the frequent problems with disambiguation more successfully. Terms often have many meanings. By, for example, giving every word the sense that is most consistent with the context in which it is used, word mapping has made it easier to avoid these errors (Casillo *et al.*, 2020). To successfully apply text analysis, a range of steps, which include text cleaning, tokenisation, part-of-speech (POS) tagging, named entity recognition (NER), sentiment analysis and dependency parsing, need to be performed. To perform the mentioned step, access is needed to NLP libraries and tools, such as the NLTK in Python, SpaCy or Stanford CoreNLP, which can perform these functions (Basha & Fathima, 2023).

According to Yates and Frank (2021), NLP tools have the additional advantage of aiding the field of social work to understand the needs and experiences of service users. Sentiment analysis can assist in identifying individuals who may be considered at risk due to their sentiment choices. Furthermore, NLP tools can enable social work researchers to identify themes and topics in large corpora of textual data, which can lead to tailored interventions and services.

2.5.3. Data pre-processing

Data pre-processing refers to cleaning and transforming raw textual data into a format that is suitable for text analysis. The purpose of data pre-processing is to prepare the text for sentiment analysis, topic modelling and NLP tasks. The lexicon is limited to the 50,000 most commonly used

words, using a unique token for out-of-vocabulary terms to reduce the dimensionality of the data and improve computational efficiency (Mehanna & Mahmuddin, 2021). Chatterjee *et al.* (2019) used the lowest number of preparation steps possible in their *tweet study*, reducing tweet capitalisation and deleting Twitter handles and hashtags. A few instances of the same punctuation that were used repeatedly were replaced. Frequently used abbreviations, such as 'FYI' were expanded into their well-known full forms ('for your information').

Critical steps in pre-processing data for a mental health chatbot would typically include the anonymisation of the text and ensuring the highest levels of confidentiality (Solanki *et al.,* 2023). Any names, phone numbers and addresses need to be removed from the text. This can be done manually or by using existing software. Techniques such as stop-word removal and stemming can also be used to improve data quality. Furthermore, the chatbot may want to use NER to identify important concepts and topics discussed by users, as well as group similar messages together for a more efficient response.

2.5.4. Input embedding

According to Chatterjee *et al.* (2019), the semantic word representations of the input words could be obtained in a variety of ways for each word in the input speech. The authors used GloVe (Yang *et al.,* 2015), Word2Vec (Mikolov *et al.*, 2013), and FastText (Joulin *et al.*, 2016). The authors utilised Sentiment-Specific Word Embedding (SSWE) (Tang *et al.*, 2014) to support the sentiment representations. SSWE seeks to encode sentiment information in the continuous representation of words. The authors trained a basic Long-Short-Term Memory (LSTM) model utilising each of the embeddings to examine the efficacy of semantic embeddings for emotion identification, which is particularly important in psychosocial professions. The concept of input embedding is crucial when developing a counselling chatbot framework to allow sensitive understanding and responses to emotions. LSTM models can capture long-term dependencies that are present in the input sequence, which assists in achieving the desired result. To assess the efficacy of various semantic embeddings, the researchers performed cross-validation. The results revealed that GloVe provides the highest macro-F1 score. The authors noticed distinct behaviours between GloVe and SSWE for 'depression' and ':'(' SSWE embeddings show a high cosine similarity; however, GloVe gave a low score although the two words have a comparable mood. SSWE correctly assigned a low score for the 'happy' and 'sad' pair, while GloVe produced a respectably high score. However, words such as 'best' and 'excellent' that are semantically related have a low cosine similarity to SSWE but a high GloVe score. Chatterjee *et al.* (2019) chose GloVe and SSWE as their embedding for semantic and sentiment LSTM based on these discoveries and their motivation for merging sentiment and semantic characteristics.

It is critical to consider the use of semantic embeddings when developing a chatbot for interpreting and understanding user input. Semantic embeddings such as GloVe, Word2Vec, and FastText

can be used to represent the meaning of words and their relationships, allowing the chatbot to understand user queries and respond appropriately. This suggests that using sentiment-specific embeddings such as SSWE can improve the chatbot's ability to detect emotions in user input, which can enable it to provide more personalised responses. This may help direct the development of chatbots that are better equipped to understand and respond to their users' emotions, resulting in a more engaging and effective conversational experience.

2.5.5. Natural Language Processing (NLP)

The study of the degree to which computer systems can comprehend and manipulate natural language in text or speech is known as NLP (Khurana *et al.,* 2022). Information on the comprehension and use of human language is gathered to develop the right methods for computer systems to manage human language and perform a variety of activities. The majority of NLP methods rely on ML. They are made up of natural language generation (NLG) and natural language understanding (NLU), which develop the task of text comprehension (Jung, 2019; Khurana *et al.*, 2022).

NLP is particularly crucial for chatbots since it controls how the chatbot will comprehend and interpret the text input. The ideal chatbot should converse with the user in such a way that the user is unaware that they are conversing with a machine. Through ML and a wealth of conversational data, the chatbot aims to understand the subtleties of human language. NLP enables the chatbot to comprehend language, sentiment, and intent when interpreting text data. In its most basic form, NLP analyses words, phrases, and sentences to determine their meaning and context before reducing them to simpler, more manageable sections. This is primarily achieved by a variety of NLP features, which include text tokenisation (when text is divided up into individual words, phrases or sentences), POS tagging (giving each word in a sentence a grammatical label), NER (the process of recognising and extracting significant named entities, such as individuals, locations, and businesses) and sentiment analysis (examining the text's emotional undertone or sentiment) (Abdul-Kader & Woods, 2015; Vijayaraghavan *et al.*, 2020). Statistical models and ML algorithms that have been trained on an abundance of text data are utilised in performing these tasks. For language analysis and prediction, Sun *et al.* (2022) note that these models employ patterns and rules discovered from the data (which are discussed in more detail in point 2.5.7.).

There are, however, limitations to NLP models, which could have adverse effects on mental health and counselling settings or contexts. Limitations that could limit the effectiveness of a counselling chatbot include not understanding sarcasm, irony, and idiomatic expressions, which are often found in counselling sessions. A concerning limitation, according to Sun *et al.* (2019), which could have far-reaching consequences in mental health services and social work services exists

surrounding the bias displayed by NLP models. Bias can be the result of the quality and quantity of data that the NLP model is trained on.

2.5.6. Naïve Bayes Method

An algorithm is a set of instructions and rules that are supplied as a specification to achieve a particular objective. Typically, algorithms process inputs and produce outputs after a certain number of finite steps, during which they cycle through a variety of states (Hurbans, 2020). There are various text classification methods, with the Naïve Bayes algorithm being one of the most frequently used (Hassan *et al.,* 2022). Using Bayes' theorem, conditional probability can be calculated statistically. Conditional probability is the likelihood that a result will occur given that a comparable result has previously occurred in the past. The Bayes' theorem provides a way for updating probabilities on existing predictions or hypotheses in the presence of new or additional evidence.

The Bayes' theorem is well-liked among scholars for text categorisations. The Bayes' theorem is a fundamental principle in probability theory that provides a method for updating the probability of a hypothesis or prediction based on new or additional evidence. It is widely used in various fields, including text categorisation, where it has been demonstrated to be a powerful instrument for creating classification models (Ball & Rague, 2022).

In the context of text categorisation, the Naïve Bayes algorithm is a popular model-based classification technique that employs Bayes' theorem to calculate the propensity of a document belonging to a specific category. It performs by first learning the probability distribution of words in each category from a training dataset before employing this information to classify new documents. As it is highly effective and simple to apply, the Naïve Bayes classification method is applicable in a wide range of situations. Each feature of the algorithm that is classified using this Bayesian approach to probability learning is independent of the value of other features. Each feature or attribute of the data is assumed to be conditionally independent of all other features, given the class label in Bayesian classifiers. That is, the presence or absence of one feature is independent of the presence or absence of any other feature (Bonaccorso, 2017).

The Naïve Bayesian Chatbot is a basic classifier for intent classification that applies the Bayesian theorem under independent assumptions (Smys & Haoxiang, 2021). Data that has been divided into numerous classes during training is necessary for classification algorithms. The model construction process for Naïve Bayes is reported to be fairly quick, as the method can be either adaptive or intelligent and can be a function that also meets the specific needs of each user. As a result, it is widely or extensively utilised to develop chatbots. The goal of applying a Bayesian method to data models is to identify the dialogue act sequence that has the highest likelihood of

occurring; this is done by calculating the probability of each conceivable sequence of dialogue acts as data that could represent a sentence or utterance (Cahn, 2017; Setyawan *et al.*, 2018).

The Naïve Bayes method has the benefit of requiring less training data to generate the parameter estimates required for classification, which is an advantage when designing chatbots. Only some of a variable's values within a class are required to determine the classification, as it is assumed to be an independent variable; the entire covariance matrix is not required (Hardeniya *et al.*, 2016). The Bayes' theorem formula is displayed as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

The Bayes' theorem could be applied to the development of a counselling chatbot that assists users in dealing with anxiety (as an example). As illustrated by Kumar and Kumar (2022), the developer could use Bayes' theorem to estimate the likelihood that a user will experience anxiety based on certain events or features:

- A = the event that the user is likely to experience anxiety; and
- B = the event that the user exhibits certain behaviours or symptoms, such as talking about feeling nervous or avoiding social situations.

Using Bayes' theorem, one can express this as:

P(A|B) = (P(B|A) * P(A)) / P(B)

- P(A|B) = probability that the user is likely to experience anxiety given that they exhibit certain behaviours or symptoms;
- P(B|A) = probability that the user exhibits certain behaviours or symptoms given that they are likely to experience anxiety;
- P(A) = prior probability that the user is likely to experience anxiety (based on demographic data or previous conversations); and
- P(B) = probability that the user exhibits certain behaviours or symptoms.

To use this formula in a counselling chatbot, one would use the user's conversation history and demographic data to estimate the prior probability of anxiety (P(A)). NLP techniques would also be used to analyse the user's language and identify specific behaviours or symptoms (B). The conditional probability of exhibiting those behaviours or symptoms given anxiety (P(B|A)) could then be estimated, for example, by training an ML model on a dataset of conversations between anxious individuals and mental health professionals. In this specific study, the training dataset consists of live chat transcriptions.

Using Bayes' theorem, one can combine these estimates to calculate the probability that the user will experience anxiety based on the behaviours or symptoms displayed (P(A|B)). The counselling

chatbot can then use this information to provide personalised support and recommendations to the user.

The Bayes' theorem can be quantified to develop an algorithm for a counselling chatbot. A summary of the steps involved at a general level is as follows:

1. Define the events of interest: In the example of anxiety, the events of interest could be ‘the user is likely to experience anxiety’ and ‘the user exhibits certain behaviours or symptoms’;
2. Gather data: Gather the user's conversation history and demographic information. Use NLP techniques to identify specific behaviours or symptoms displayed by the user;
3. Calculate prior probability: Based on demographic information or previous conversations, calculate the prior probability of anxiety (P(A));
4. Train an ML model: To estimate the conditional probability of exhibiting certain behaviours or symptoms given anxiety (P(B|A), train an ML model on a dataset of conversations between anxious individuals and mental health professionals;
5. Apply Bayes' theorem: Using Bayes' theorem, combine the estimates from Steps 3 and 4 to calculate the probability that the user will experience anxiety based on the behaviours or symptoms displayed (P(A|B)); and
6. Provide support: The chatbot can provide personalised support and recommendations to the user based on the probability calculated in Step 5.

As the chatbot engages in conversations with users and gathers more data, it can update its prior probability of a user experiencing anxiety, leading to more accurate predictions. For example, if the chatbot encounters a user who exhibits certain behaviours but does not report anxiety, the chatbot can update its prior probability of the user experiencing anxiety to a lower value.

Based on the estimated probability of a user experiencing anxiety, the chatbot can offer relevant resources or advice to help the user cope with their symptoms. This can include certain techniques or referrals to mental health professionals (Tebenkov & Prokhorov, 2021).

2.5.7. Machine learning (ML)

According to Suta *et al.* (2020), modern automated chatbots with advanced capabilities typically combine ML and NLP within an AI domain. By employing NLP to extract the information from user input, chatbots that use ML approaches instead of pattern matching can learn from interactions. They do not require a prepared response for every potential user input and instead consider the entire conversation context, not just the current turn. Usually, they require a large training set and finding one might be problematic because the datasets that are currently accessible may not be sufficient (Adamopoulo & Moussiades, 2020). It is crucial, according to Chatterjee *et al.* (2019), to use big data and crowd intelligence to solve the problem of emotion recognition, as emotions can be represented in a variety of ways.

Woebot, as referred to in Chapter 1, draws upon ML to successfully deliver mental health services to its users. The success of Woebot's ML capabilities and outputs is evident in a study by Prochaska *et al.* (2021) on a sample of 101 users, which found that from pre- to post-treatment:

- Confidence in resisting urges to use substances increased significantly (mean score change +16.9, SD 21.4; P.001);
- Past month substance use occasions decreased (mean change 9.3, SD 14.1; P.001);
- The Alcohol Use Disorders Identification Test (AUDIT)-Concise (mean change 1.3, SD 2.6; P.001);
- The 10-item Drug Abuse Screening Test (mean change 1.2, SD 2.0; P.001);
- The Patient Health Questionnaire-9 item (mean change 2.1, SD 5.2; P =.005);
- The Generalised Anxiety Disorder-7 item (GAD-7)(mean change 2.3, SD 4.7; P =.001); and
- The Cravings scale (68.6% versus 47%).

From the standpoint of social work and an EHWP, it is critical to recognise the potential of ML and AI in providing mental health services to individuals, particularly those who may not have access to traditional face-to-face services. ML does pose the risk of dehumanising the therapeutic process and chatbots may, in certain instances, not be equipped to deal with complex issues of trauma, for example. This may require the expertise and judgment of a human clinician to ensure optimal outcomes. (Zhu *et al.*, 2022).

2.5.8. Sentiment analysis

Sentiment analysis is a subfield of NLP that examines the strength of the emotions expressed in a text. How sentiment (-ve or +ve) is expressed in textual information depends on how people think. Sentiment analysis makes it feasible to examine the emotions or sentiments of textual communication through dictionary-based sentiment analysis or ML techniques. It is, therefore, crucial to identify aspect extraction and sentiment polarity from written texts to design any functioning chatbot (Bergman & Humble, 2022; Pradhan *et al.*, 2022; Keenan, 2019).

Sentiment analysis, according to Nandwani and Verma (2021), is not yet a perfect science but a crucial component of a counselling chatbot, as this allows the chatbot to provide personalised and empathic responses to the user. Sentiment analysis has numerous applications and can be performed using a variety of approaches. These approaches are classified into three categories, namely lexicon-based, ML-based, and DL-based. Each of these approaches has its variety of advantages and disadvantages. Despite several sentiment and emotion recognition systems, researchers encounter substantial hurdles, such as analysing context, mocking, statements conveying multiple emotions, Web slang propagation, and lexical and syntactical ambiguity.

According to Mehta and Pandya (2020), the lexicon-based sentiment analysis method is a basic, easily understood method that uses a lexicon or dictionary of terms with positive, negative or

neutral sentiment scores. The words in a text are then rated based on their sentiment value, with the overall sentiment of the text calculated by the sum of all the scores. The two types of lexicon-based approaches, namely dictionary-based and corpus-based approaches, are widely used. In this approach, a vocabulary of sentiment words is manually developed by a human specialist in the dictionary-based technique. The corpus-based technique, on the other hand, employs a corpus of text to construct a vocabulary of sentiment terms automatically. The lexicon-based technique is a basic and easy-to-implement method for rapidly and simply determining the sentiment of a text. However, it is crucial to highlight that the lexicon-based approach is not always accurate, as finding sentiment words for all possible contexts can be challenging. Furthermore, the lexicon-based method is incapable of accounting for human linguistic characteristics such as sarcasm and irony.

The classification in the ML approach is initially conducted using two alternative assemblies of the document. These include both trained and test data. This is known as involuntary categorisation. Text is retrieved from the characteristics and classified as supervised or unsupervised. Supervised learning draws on classifiers, which are categorised into linear classifiers, decision tree classifiers, rule-based classifiers and probabilistic classifiers (Coria *et al.,* 2021).

Sentiment analysis is essential when creating a chatbot, as it enables the bot to comprehend the emotional tone of the user's input. This information can then be used to improve the chatbot's response. If a user portrays a negative emotion, the chatbot may suggest that they be transferred to a live agent. If the user portrays a positive emotion, the chatbot may offer additional information or services. Sentiment analysis has the additional benefit of monitoring the changes in the user's emotional states over time, which will allow the chatbot to provide the required support and complete the appropriate referrals (Bhangdia *et al.*, 2022; El-Ansari & Beni-Hssane, 2023), which are critical components of effective EHWP services.

2.5.9. Artificial Neural Networks (ANN)

In addition to the traditional rule-based approach that was previously used in chatbot development and other simple ML algorithms, more advanced ideas and methods are now being deployed to model and train chatbot systems. According to Dhyani and Kumar (2021), this includes DL methods such as Deep Neural Network (DNN) and Deep Reinforcement Learning (DRL). Artificial neural networks (ANN) are computer models inspired by biological systems that are built from various artificial neurons which are coupled together by coefficients (weights) to construct the neural architecture as depicted in Figure 12. ANN is capable of self-learning and self-adaptation. It can offer a batch of input and output neurons that correspond in advance, analyse the relationships and rules that exist between the neurons, and then create a complex non-linear function using these relationships and rules. The term ‘training’ refers to this process of learning

and analysis. The strength of each synaptic connection between a neuron (input, hidden, and output neurons) is indicated by a connection weight (Lee *et al.*, 2022).

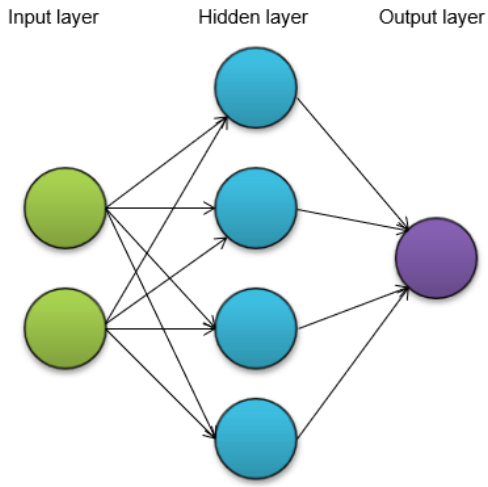


Figure 12. Example of a simple ANN

ANNs are a powerful tool; they improve the capabilities and performance of intelligent chatbot systems in that they facilitate intent recognition, sentiment analysis, NLP, and contextual understanding, which are critical elements for a counselling chatbot. ANNs are important to a counselling chatbot, as they enable the chatbot to learn and adapt its responses based on its interactions with users. In the case of a counselling chatbot, a neural network could be trained on large amounts of text data from previous counselling sessions (live chat transcripts in this study), enabling it to learn from previous interactions and provide more personalised and effective responses to users.

For example, the neural network could be trained to recognise patterns in the types of problems users seek help with and provide appropriate responses based on this analysis. It could also learn to adapt its responses based on user feedback, improving its ability to provide effective assistance over time. Overall, the use of ANNs can aid in improving the effectiveness of a counselling chatbot and provide more personalised and helpful support to users (Shum *et al.*, 2018).

2.5.10. Integration of research methodology in the facilitation of chatbot design

The research methodology is discussed comprehensively in Chapter 1. To complete the necessary chatbot design methods, namely text analysis, data pre-processing, input embedding, NLP, the Naïve Bayes classifier, ML, sentiment analysis and ANNs, a unique approach to social work research methodology was designed. The design consists of a multi-phased, mixed-method approach in the form of a qualitative first and a quantitative second data analysis.

Live chat transcripts for a period of six months, consisting of 50 or more interactions between human counsellors and clients of Lyra SA that took place on the platform *Cubit*, were subjected to thematic and textual analysis. This design has resulted in the acquisition of big data sets that are essential for informing the aforementioned components of an intelligent online counselling framework for a social work EHWP.

2.6. THE USE OF BIG DATA IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE (AI) CHATBOTS

Big Data is a notion that has become increasingly popular over the past several years and is highly relevant in today's society. It is often highlighted as a crucial asset for improving efficiency, innovation, and consumer relationships. Smart cities, manufacturing, retail, finance, software development, the environment, and digital media are just a few of the industries that stand to gain from the collection, storage, processing, and analysis of big data. These industries will be able to take advantage of state-of-the-art data-driven workflows and significantly enhanced decision-making processes (Santos & Costa, 2020). Porter *et al.* (2020) maintain that the adoption of big data in many social sciences has been slow. Big data's transformative potential in social sciences has been compared to how the invention of the telescope revolutionised the study of space; however, some areas of social science have been slow to embrace big data. According to Zetino and Mendoza (2019), the role of social work in shaping this revolutionary shift is critical. The authors argue that data-driven outcomes should be integrated into social work practice to improve care and reduce costs.

The definitions of big data are vast. For instance, some scholars such as Santos and Costa (2020:12) characterise big data as "enormous datasets" that can be unstructured, semi-structured or both. The use of terabytes or exabytes to characterise big data is an imprecise science since the scale of large data is continually changing with technological advancements. Regardless, it is acknowledged that big data are many, intricate, and enormous in scope, necessitating specific analytic tools and technologies to gather, analyse, and quickly disclose insights. The three V's-volume, velocity, and variety - are frequently used to describe the properties of big data. Although theorists accept that the relative size changes across multiple disciplines, volume refers to the bulk scale of data. Big data is now being produced at a rate that has never been seen before and is continually rising. Velocity is the quick creation of information or timeliness. Variety refers to the numerous modalities or sorts of structured, semi-structured, and unstructured data, such as audio, text, photos, numbers, and other types of data. Big data can come in many different forms, such as social media activity logs, corporate and operational data, web and mobile analytics, streaming data, commercial and government data, and sensor data from the IoT.

Other uses for big data include automated data aggregation and mining, sentiment analysis, opinion mining, computer-assisted content analysis, visualisation of massive data sets, ML, sentiment analysis, and computer-assisted content analysis (Mills, 2019; Pencheva *et al.*, 2020). Big data analytics, according to Sankar (2018), is the process of analysing vast amounts of data to uncover new patterns and undiscovered relationships, among others. The analytical results can be employed in a variety of disciplines and result in a more efficient use of the information obtained. Traditional methods of data analysis are insufficient and specialised tools are needed to analyse big data.

The scientific community and real-world phenomena have both been productively transformed by AI and big data. Big data appears to be reviving the notion that AI has applications outside of academia (Hussain & Manhas, 2016). Khan and Alotaibi (2020) note that big data analytics offers a variety of alternatives, with specific reference to client care and clinical support in the healthcare industry. The use of big data in healthcare typically refers to the management of large, complicated, and difficult-to-manage electronic health datasets using standard technology, software, tools, and procedures. Clinical information obtained from clinicians and other data on administrative processes constitute big data in healthcare. Big data and AI are often related concepts that complement one another as catalysts for the current digital transformation.

According to Lommatzsch and Katins (2019), constructing and training a chatbot usually necessitates substantial amounts of training data from the target field. This method of data collection, when combined with NLP, aids in the understanding of human communication. It is simple to recognise patterns and tailor messaging by studying the language users use, alongside their demographics. Learning algorithms have been created specifically to benefit from exceedingly big data sets. The great majority of examples in these data sets are frequently left unstructured. For instance, in Yarowsky's (1995) work on word-sense disambiguation, instances of the word ‘plant’ that might be used to denote a factory *or* flora are not tagged in the data set. The process of determining which sense was meant in a statement, however, may be accomplished with an accuracy of over 96% by using the right learning algorithms and sufficiently big data sets (Russel & Norvig, 2022). The novel insights can be used by service providers that deploy chatbots to speak in the language of their audience in emails, advertisements, and even through the chatbot itself.

In-depth analysis is where big data's true value resides. This relates to developing specific business intelligence tools regarding chatbots. A Chabot’s primary algorithm is based on pattern recognition. In essence, the bot cannot comprehend what the user is saying or asking. More advanced chatbots can analyse a user's speech to identify word groups in context and assign a certain emotion. This process, known as sentiment analysis, which is based on big data, enables service providers to identify customer emotions early and address root causes (Brooke, 2017).

The use of big data does not cease once an AI chatbot has been developed and deployed. Sinha and Dubey (2022) believe that sustainability combined with AI and big data will provide a company or organisation with a competitive advantage. By utilising big data generated by AI, a company could base its strategies on sustainability and adjust data-driven information to improve sustainability, which will result in cost savings and other advantages.

Big data is transforming the healthcare industry and how services are delivered. Chatbots require big data to be trained to serve the target population. Social work has a significant role to play in adopting big data, including in the development of AI counselling chatbots. The incorporation of data-based outcomes into social work, specifically an EHWP, can improve care and reduce costs. According to Zetino and Mendoza (2019), in-depth analysis of data is where the true value of big data is situated, as it relates directly to specific business intelligence tools, such as mental health chatbots for an EHWP.

2.7. THE USE OF LINGUISTICS IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE (AI) CHATBOTS

Chatbots identify particular keywords to deliver the intended response. They give importance to what users enter and react appropriately. The overarching concept within this project would ideally mimic a social work interview, drawing on insights from social work and psychological theory within an EHWP context, as conceptualised by Archard (2021). The bot chooses an acceptable response for the user by utilising algorithms with the aid of AI technology and customised keyword lists. Artificial Intelligence Markup Language (AIML), a sophisticated markup language evolved from Extensible Markup Language (XML), is used to create distinctive chatbots. Since AIML-based chatbots are lightweight (a software application that is intended to consume less memory) and relatively simple to configure, it is simpler to create more AI applications. These methods analyse real language and use a structured framework to train the chatbot (Gupta *et al.*, 2020).

Natural language should be understood and used by chatbots to communicate. The findings of a study conducted by Chaves and Gerosa (2022) confirm the impact of using register-specific language (how humans adapt their language use depending on the interactional situation) on user perceptions of their interactions with chatbots. Their study reinforced the importance of considering register when designing chatbot language. Chaves and Gerosa (2022) demonstrated that using language that is register-specific within a specific context has a significant impact on users' perceptions of interactions with chatbots. Carefully planning a chatbot's use of language is one way to improve its impersonation of a social role. When a chatbot employs unexpected levels of (in)formality or incoherent language style, the interaction can be frustrating or awkward.

Modern, sophisticated automated chatbots typically combine ML and NLP in the domain of AI. Researchers (Adamopoulou & Moussiades, 2020; Suta *et al.*, 2020) have determined that

chatbots' anatomy for information-seeking typically consists of three parts, namely NLU to classify the user's intent, DM to ascertain the user's intent, and NLG to produce a response in natural language. A chatbot's knowledge base can be improved by using language tricks to make it appear more human. These methodologies mimic conversational behaviour by simulating stereotypical responses, typographical errors, typing style, the presence of a personality, as well as irrational responses.

2.8. THE USE OF AI CHATBOTS WITHIN A SOCIAL WORK AND AN EHWP CONTEXT

In the late 19th and early 20th centuries, occupational social work was recognised as one of the core areas of social work interventions, requiring a domain-specific knowledge and skillset. It evolved from the principles and background of generic social work practice into a specialised field, subsequently forming part of EHWPs. Occupational social workers are either employed by the organisation or deployed as part of outsourced services to provide counselling and interventions to employees and their families as part of EHWPs. EHWPs are described as an accumulation of systematically designed programmes that aim to address issues affecting employee productivity. These programmes emphasise employees' physical and mental wellbeing, which allows them to focus on their work, resulting in increased productivity (Chen *et al.,* 2021; Sánchez *et al.,* 2020; Terblanche, 2021).

Research on AI's application in the helping professions is still in development, especially in social work and EHWPs. However, there is significant potential for scientific research with a focus on the use of chatbots as an application of AI (Terblanche, 2020). According to Safodien (2021), social work must reap the benefits and address the difficulties that the 4IR brings. To this end, it might be beneficial to examine how the rapid technological advances of this era affect the social work profession and its progress. As per Güzel (2021), technology, digitalisation, AI, and design should be incorporated into social work curricula, and social policies and social work practices should be updated to reflect recent advancements.

AI4SG is a research area regarding the use and advancement of AI to address societal problems and enhance societal wellbeing. AI4SG is formally defined as the design, development, and deployment of AI systems in ways that i) enable socially preferable or environmentally sustainable developments; ii) prevent, mitigate, and/or resolve problems that harm human life and/or the wellbeing of the natural world, and iii) do not introduce new types of harm or amplify current disparities and inequities (Cowls *et al.*, 2021). The research community has paid attention to AI4SG over the past 10 years as a result of its many successful applications. To address societal issues and enhance societal wellbeing, researchers are developing various AI methodologies and technologies. Several AI4SG applications have recently proven effective, for example the application offering HIV prevention awareness to homeless adolescents (Rice *et al.,* 2021).

Large organisations are increasing their investment in AI4SG programs (Holzmeyer, 2021), as a result of the recognition that AI4SG has received from both within and outside of the academic community. It is commonly acknowledged that to solve AI4SG challenges, researchers, governments, businesses, and non-profit organisations must collaborate. It is reasonable to claim that models of how humans explain decisions and behaviour to each other are a suitable place to start when analysing the issue of designing and implementing intelligent entities that are truly capable of offering services to people.

According to Goldkind (2021), social work must address the implications of interacting with embodied algorithms, develop the capacity to assess data sources and broaden the definition of collaboration to consider the inevitable need to collaborate with algorithmic actors such as chatbots and other technologically advanced tools. Social work may need to broaden its fundamental concepts of both ‘person’ and ‘environment’ to engage with augmented and virtual worlds and have a conversation with computer agents, such as robots and algorithms. Although AI-based solutions are not able to resolve all issues, Cowls *et al.* (2021) believe that they can assist with the most pressing social concerns that society is currently facing. AI technologies have the potential to produce socially beneficial results at a scale and efficiency never before possible if they are designed appropriately, for instance, as suggested in Figure 13.

Roestenburg's (2021) research mentions that social workers, while not studied specifically in the context of EHWPs, held a positive outlook towards IT technology and expressed a willingness to utilise it if it was introduced to them. According to the results, the social service sector is ready to embrace IT technology and the market is ready to accept the technology. This can be achieved by incorporating the following recommendations that emerged from the study:

- A thorough understanding of market needs must be obtained to develop an appropriate system that reflects the needs and practices of the target market segment;
- Any technology development efforts should consider the constantly changing state of information needs and demands on the sector;
- Any product introduction should be carefully designed because social workers in the social services sector frequently exhibit characteristics similar to their counterparts in the healthcare sector in terms of their proclivity to override organisational priorities and demands to meet their clients' service demands and needs;
- There is no single IT system that can meet all of the needs of the organisation or industry for which it was designed. Technology adoption and implementation appear to be a continuous, time-consuming process;
- Adoption and implementation of technology are strongly linked to changes and improvements in an organisation's operational processes; and
- Adoption of IT should be part of a long-term strategy.

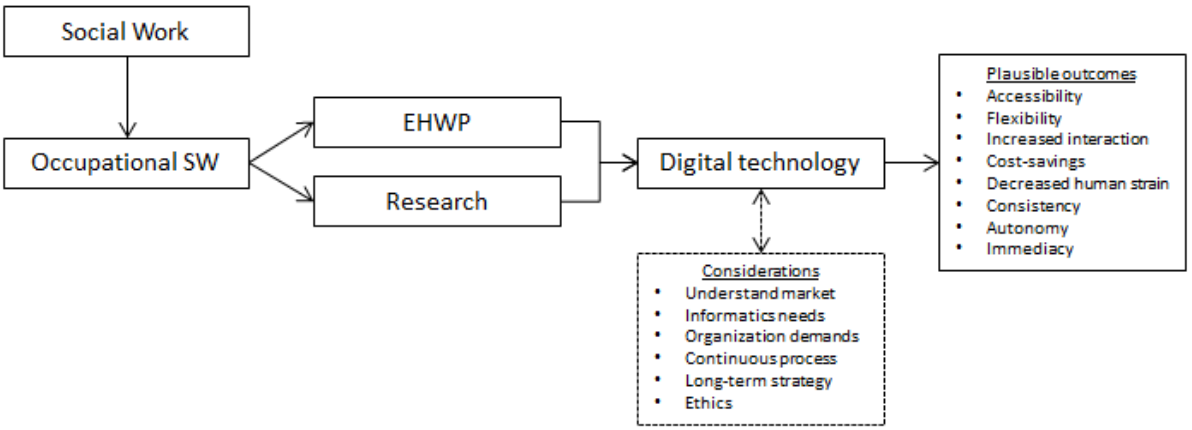


Figure 13. Integration of IT into social work and EHWPs

Although Morrissey (2021) does not refer to a specific chatbot in social work, they note that chatbots are currently being used to assist social workers with everyday activities, which most notably include directing service users in need to specific resources that they may require. According to Sage and Singer (2022), current AI algorithms in social work are unregulated but are being used in the United States of America to determine appropriate jail sentences using risk scores based on historical data. Algorithms, although controversial, are also used in child welfare to determine the risk profiles of children and families. The authors continue to emphasise the potential benefits of chatbots in social work as well as the capacity of technology to efficiently assist clients without the need to access already scarce human resources (Sage & Singer, 2022).

Although no EHWP counselling chatbot has been developed and produced, Inkster *et al.* (2018) have studied employees from a certain company that installed and used Wysa on their devices Wysa is intended to assist individuals who are currently struggling with low mood, stress or anxiety or those who aim to enhance their emotional resilience. The study shows that the average mood improvement (difference in pre- and post-self-reported depression scores) between the test groups (high versus low users; n=108 and n=21, respectively) revealed that the high users had significantly higher average improvement (mean 5.84 [SD 6.66]) compared to the low users (mean 3.52 [SD 6.15]); Mann-Whitney P=.03; and a moderate effect size of 0.63. The results are a positive predictor of what a counselling chatbot within an EHWP context could achieve.

By keeping up with technological advancements and the 4IR, AI will influence how psychosocial professionals work in the future, as it has the potential to significantly alter how wellness is organised in any business or organisation. Baldwin (2019), however, predicts that it will take approximately 50 years for AI to match top-level human performance in the four social skills that are important in the workplace, namely social and emotional reasoning, coordination with many people, acting in emotionally appropriate ways, and social and emotional sensing. When studying the use of AI in social work, specifically within an EHWP context, Srinivasan (2021) concludes

that AI has demonstrated an ability to transform human resource services, especially within the more complex domain of employee health and wellness. The data gathered by the author indicates that employees are starting to seek their employer's mental health resources more often than previously. Conversations regarding mental health are increasingly being initiated via AI chatbots. This is especially evident in the younger generation of employees.

The use of chatbots as an application of AI within the context of social work and an EHWP is a developing area with significant potential. AI methods and technologies are being developed within the AI4SG research arena to address social problems and improve the wellbeing of society. The introduction and implementation of technologies are closely linked to changes and improvements in an organisation's operations. AI-based solutions may not solve all problems but they can help with the most pressing social issues currently facing society. Chatbots have the potential to efficiently assist service users without having to rely on already scarce human resources, and the market is ready to embrace the technology. However, there is a need to address the implications of interacting with embodied algorithms, develop the ability to evaluate data sources and broaden the definition of collaboration to include the inevitable need to work with algorithmic actors such as chatbots and other technologically advanced tools.

2.9. CLINICAL THEORETICAL FRAMEWORK

The selected theoretical framework and therapeutic approach preferred by the hosting company is the solution-focused brief therapy (SFBT) paradigm (Molnar & de Shazer*,* 1986). For this reason, all data utilised for analysis and development of the proposed framework draw upon the principles of SFBT. As a result of all the data and training data being related to SFBT, it should be obvious that the proposed product will also function within the SFBT paradigm.

The strengths-based approach of SFBT is effective for crisis care across treatment sites, presenting problems and demographics. In the SFBT paradigm, a mental healthcare practitioner collaborates with a client to critically evaluate their subjective perceptions of reality to construct a new or modified version of reality as a catalyst for change in the SFBT counselling process. Since SFBT is a time-limited counselling modality, the primary focus is on identifying positive solutions to difficult behavioural problems rather than on causes, deficiencies or limitations (Johnson *et al.*, 2015).

The proposed product corpus should integrate the most prominent SFBT principles, which include (De Jong, 2015; Finlayson *et al.*, 2023):

- Clients who are competent to co-construct goals;
- Clients who are the authorities on their own lives and the significance of their experiences;
- That there may not always be a link between the problem and the solution;
- Clients who must take different actions for change to occur; and

- Change that starts with small changes, after which larger changes could follow.

Further important considerations in SFBT include the formulation of questions that are goal-orientated. After the client briefly describes the problem or situation, the co-construction of goals ensues by asking:

- When the problem is resolved, what will be different?; and
- What would have to be different by the end of the session for you to say our time together was useful?

The miracle question, which is probably one of the most distinct features of SFBT, will be crucial to the proposed product's probing and solution-generating capabilities. This is because this question serves as an opener for a series of follow-up questions and responses. Further questions that need to be considered in the proposed product's response would include, but are not limited to:

- Exception-finding questions;
- Scaling questions;
- Coping questions; and
- 'What's better' questions.

The question arises as to whether delivering SFBT is possible by using textual communication. Wulf (2020) has explored how SFBT practitioners incorporated linguistic focus into their practice. The author states that textual dialogues differ significantly from in-person therapeutic dialogues. The fact that the practitioner and client cannot see each other presents certain challenges in terms of synchronous and asynchronous communication, such as with a pause. Lexical choices in the practitioner's questions, responses, and summaries are therefore critical. Wulf (2020) concludes that the limitations of SFBT via textual communication can be resolved by the correct use of wh-questions, such as 'what,' 'when,' 'where,' 'why,' and 'who' questions. These assumptions will be investigated through the textual analysis of live chat transcripts within this study. The outcomes of successful engagements will provide the evidence that is required to create the conceptual counselling framework to inform an AI algorithm.

2.10. ETHICAL CONSIDERATIONS IN THE DEPLOYMENT OF AI IN A SOCIAL WORK EHWP

Despite the potential benefits of technology in mental healthcare services, its limitations must be acknowledged. The negative aspects of chatbots were most visible in international media coverage of Woebot in December 2018, when the BBC demonstrated that Woebot failed to detect difficult and dangerous situations. This occurred because of a boy end-user who communicated the message, "I'm being forced to have sex, and I'm only 12 years old". Woebot responded in an

extremely inappropriate manner: “Sorry you’re going through this, but it also shows me how much you care about connection, and that’s really kind of beautiful”. Similarly, an examination of chatbot responses to suicide emergencies revealed that they are often limited and inappropriate (Vaidyam *et al.*, 2019).

The unintended consequences of such technologies in resource-constrained mental health ecosystems appear to be underappreciated. According to Kretzschmar *et al.* (2019), there has been no independently conducted longitudinal research on the effects of automated CBT over time (at the time of writing), and most digital mental health tools on the market have not been evaluated through an ethical appraisal.

To mitigate the ethical risks posed by therapeutic AI agents, Vilaza and McCashin (2021) recommend embedding the ethical principles of beneficence, non-maleficence, autonomy, justice, and explicability in the design of any AI therapeutic framework. These ethical recommendations resonate with those of the SACSSP’s general ethical standards for social workers and the National Association of Social Workers’ (NASW) code of ethics. After analysing 233,914 tweets on ChatGPT, Taecharungroj (2023) concludes that it is critical to accelerate debates and address the discord between technological progress and ethical concerns. This function can, from the onset of the implementation of the proposed product, be monitored by qualified clinicians.

2.11. RECOMMENDATIONS

For the field and science of social work to be relevant in a rapidly changing society, it is recommended that social work as a profession and as an academic field heed the demand for digitalisation and conformity to automation, IoT, and the 4IR. In-depth research and cross-disciplinary cooperation are required for this novel concept to become a reality.

Digital technologies and automation in social work have the potential to significantly improve the profession's ability to address complex social issues and promote social justice. Digital tools and IoT can improve communication, data management, and service delivery while also increasing access to resources for underserved populations. To ensure that the benefits of these technologies are distributed equitably, social work researchers must engage in ongoing research, collaboration, and critical reflection to identify and address the potential ethical, legal, and social implications of these developments.

Furthermore, to prepare the next generation of social workers for the demands of the 4IR, social work education and training must incorporate digital competencies such as data analysis, programming, and cybersecurity. Social work can fulfil a critical role in promoting positive psychosocial change and advancing the wellbeing of individuals, families, and communities in the 21st century by embracing digitisation and cross-disciplinary collaboration.

It is essential to establish collaboration between mental healthcare professionals and IT professionals to realise the aforementioned recommendations. It is also necessary to conduct a comprehensive analysis of big data, develop a chatbot corpus framework, conduct rigorous testing and refinement of the framework, and share the research outcomes. These steps are critical to the successful implementation of the proposed recommendations.

The research methodology proposed to inform this study aims to achieve the aforementioned by employing a multi-phased, mixed-method research approach that will analyse big data utilising conventional and AI analysis software. The approach proceeds to develop the proposed framework, which will then be subjected to scrutiny by mental healthcare practitioners as well as IT experts employing the Delphi method. The recommendations will be implemented to the extent that consensus regarding the proposed framework is reached. The findings and framework will then be made available for further research, development, and refinement.

It is important to note that this study does not propose in any way that social workers or mental healthcare practitioners can or should be replaced by AI in any form of practice or domain of service. However, for social work to adapt to and contribute to the demands of 4IR and the rapidly changing digital landscape, it is recommended that social work as a profession focus on and implement the following elements:

- Embrace digital data management tools: Data can be managed by utilising digital tools, such as electronic health records and other cloud-based systems. Digital data management tools have the potential to increase efficiency, reduce errors, reduce costs, and provide more accurate data for case management;
- Implement digital interventions: Chatbots and self-help apps provide a platform to conveniently access mental health services, especially for service users who may have limited access to traditional services;
- Develop digital competencies: Competencies in data analysis, programming and cyber security should be integrated into social work practice to improve service delivery;
- Establish collaboration between mental healthcare professionals and IT professionals. This collaboration can enable social workers to develop and implement user-friendly, effective digital interventions and applications that can meet the needs of their clients;
- Incorporate ethical considerations: The ethical implications of digital technologies and automation in social work services must be considered. This includes ensuring privacy and confidentiality, avoiding bias in algorithms and data analysis, and addressing the potential limitations of chatbots and other automated tools;
- Conduct ongoing research: Social workers should conduct ongoing research and critical reflection to identify and address potential ethical, legal, and social implications of digital technologies and automation in social work practice. This includes assessing the

effectiveness of digital tools and interventions, identifying potential risks, and developing strategies to mitigate these risks; and

- User-centred design in the development of digital interventions: It is essential to involve service users in the design and testing of digital interventions to ensure that they are effective, accessible, and meet the needs of the people they are intended to serve.

2.12. CONCLUSION

As society is entering the 4IR, it seems fitting that social work adapts its approaches to prevent lagging regarding IoT, automation, and digitisation. Social work could deploy AI in various domains of science. AI has been deployed to perform certain functionalities, which are described as AI4SG. The deployed technologies are few and far apart, and largely focused on specific areas of service delivery. As the challenge has risen for social work to modernise its approach, it seems fitting that AI be introduced in the form of an AI therapeutic chatbot. EHWP services are offered to entities' employees as part of their corporate responsibility. Many employees or EHWP service users, especially the younger generation, have displayed an interest in the utilisation of virtual services and online technologies. All current technologies are, however, still being manned by qualified mental healthcare professionals through digital services such as live chat. As with human-managed text-based therapeutic services, automated services have proven to have certain limitations and benefits for both the service user and the service provider.

Using a multi-phased, mixed-method approach, this study proposes the exploration and design of a conceptual AI generative chatbot counselling framework to inform an AI algorithm, specifically in the field of social work, that focuses on the needs of users of a larger EHWP. Chatbot development is no easy feat and requires the collaboration of IT professionals as well as mental healthcare professionals. Mental healthcare modalities need to be built into the AI agent utilising various processes, which commence with the development of a framework. To develop an AI framework, analysis needs to be conducted on big data sets. The framework is then utilised to develop and train a chatbot corpus.

To develop an AI counselling chatbot framework, big data consisting of live chat transcripts, based on the principles of SFBT and the use of linguistics in text-based counselling services need to be analysed. Once the data have been analysed, the proponents of the proposed framework could be used to inform input embedding, NLP, the Naïve Bayes classifier, ANNs, and finally ML. To achieve this, live chat transcripts must be pre-processed and analysed using well-known analytical software and various existing AI technologies, which are not often employed in social work research.

Studies on the topic of the use of AI chatbots in the field of social work within an EHWP are limited. As such, the development of an AI chatbot framework, drawing on the principles of SFBT and

incorporating ethically sound practice, within this study could offer the opportunity for additional research and development on the topic. Social work, as well as other disciplines, could make use of the research outcomes to design, deploy, and develop an AI chatbot.

REFERENCES

Abd-alrazaq, A.A., Alejano, M., Alalwan, A.A., Bewick, B.M., Gardner, P. & Househ, M. 2019. An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132, art. #103978. doi:10.1016/j.ijmedinf.2019.103978

Abdul-Kader, S.A. & Woods, J. 2015. Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*, 6(7):72-80.

Adams, D. 1979. *The hitchhiker's guide to the galaxy*. New York, NY: Harmony Books.

Adams, D. 1980. *The Restaurant at the End of the Universe*. New York, NY: Harmony Books.

Adams, D. 1982. *Life, the Universe and Everything*. New York, NY: Harmony Books.

Adamopoulou, E. & Moussiades, L. 2020. *An overview of chatbot technology*. Paper delivered at IFIP International Conference on Artificial Intelligence Applications and Innovations, Cham. https://link.springer.com/chapter/10.1007/978-3-030-49186-4_31 Date of access: 18 Dec. 2022.

AFP Wire Service. 2023. LOOK: ChatGPT bot passes US law school exam. *IOL,* 28 Jan. https://www.iol.co.za/technology/software-and-internet/look-chatgpt-bot-passes-us-law-school-exam-7eade96a-fde5-401d-9536-04e18d2fd0a1 Date of Access: 28 Jan. 2023.

Ahmad, N.A., Che, M.H., Zainal, A., Abd Rauf, M.F. & Adnan, Z. 2018. Review of chatbot design techniques. *International Journal of Computer Applications,* 181(8):7-10.

Archard, P.J. 2021. The psychoanalytically-informed interview in social work research. *Journal of Social Work Practice*, 35(2):191-203.

Ashton, K. 2009. That 'internet of things' thing. *RFID Journal*, 22(7):97-114.

Ball, R. & Rague, B. 2022. *The Beginner's Guide to Data Science*. Cham: Springer International Publishing.

Baldwin, R.E. 2019. *The Globotics Upheaval: Globalization, Robotics, and the Future of Work*. New York, NY: Oxford University Press.

Bansal, H. & Khan, R. 2018. A review paper on human computer interaction. *International Journal of Advanced Research in Computer Science and Software Engineering,* 53(8), art. #8 doi:10.23956/ijarcsse.v8i4.630

Basha, S.M. & Fathima, A.S. 2023. *Natural Language Processing: Practical Approach*. Karnataka: MileStone Research Publications.

Bendig, E., Erb, B., Schulze-Thuesing, L. & Baumeister, H. 2019. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. *Verhaltenstherapie*, 29(4), art. #266. doi:10.1159/000501812

Balcombe, L. & De Leo, D. 2022. Evaluation of the Use of Digital Platforms and Digital Mental Health Interventions: Scoping Review. *International Journal of Environmental Research and Public Health*. 20(1):362.

Bergman, S. & Humble, N. 2022. Sentiment analysis for emotional navigation in written communication: What support do autistic people need? *Artificial Intelligence and Robotics*, 4(1):6-11.

Bhangdia, C., Jadhav, S., Gadgil, T., Kumari, A. & Dasari, M. 2022. Sentiment Analysis using Chatbot and Mental Health Tracker. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1):131-136.

Bridgman, L.A. & Lazar, B.S. 2019. A Systematic Literature Review of the use of Artificial Intelligence (AI) or Mobile Applications for Depression Therapy. *International Journal of Academic Health and Medical Research*, 3(7):18-24.

Brooke, S. 2017. *3 Revolutionary Ways of Using Chatbots for Big Data*. https://chatbotslife.com/3-revolutionary-ways-of-using-chatbots-for-big-data-c0a9076171c3 Date of access: 10 Jan. 2023.

Bonaccorso, G. 2017. *Machine learning algorithms*. Birmingham: Packt Publishing.

Boucher, E.M., Harake, N.R., Ward, H.E., Stoeckl, S.E., Vargas, J., Minkel, J., Parks. A.C. & Zilca, R. 2021. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Review of Medical Devices*, 18:37-49.

Cahn, J. 2017. *Chatbot: Architecture, Design, & Development*. Philadelphia: University of Pennsylvania (Thesis – PhD).

Casillo, M., Clarizia, F., D'Aniello, G., De Santo, M., Lombardi, M. & Santaniello, D. 2020. CHAT-Bot: A cultural heritage aware teller-bot for supporting touristic experiences. Pattern *Recognition Letters*, 131:234-243.

Chatterjeea, A., Guptaa, U., Chinnakotlab, M.K., Srikanthb, R., Galleyc, M. & Agrawala, P. 2019. Understanding Emotions in Text Using Deep Learning and Big Data. *Computers in Human Behavior*, 93:309-317.

Chaves, A.P. & Gerosa, M.A. 2022. *The Impact of Chatbot Linguistic Register on User Perceptions: A Replication Study*. Paper delivered at the Chatbot Research and Design: 5th International Workshop, Virtual Event. https://link.springer.com/chapter/10.1007/978-3-030-94890-0_9 Date of access: 08 Apr. 2023.

Chen, Y., Chu, H. & Wang, P. 2021. Employee assistance programs: a meta-analysis. *Journal of employment counselling*, 58:144-166.

Chivers, T. 2019. *The AI Does Not Hate You: Superintelligence, Rationality and the Race to Save the World*. London: Weidenfeld & Nicolson.

Colby, K.M., Weber, S. & Hilf, F.D. 1971. Artificial paranoia. *Artificial Intelligence*, 2(1):1-25.

Coria, M.D., Arevalo, J.C.M., Hilario-Rivas, J.L., Hilario-Cárdenas, J.R. & Prado-Juscamaita, J.I. 2021. Supervised Sentiment Analysis Algorithms. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(14):2000-2012.

Cowls, J., Tsamados, A., Taddeo, M. & Floridi, L. 2021. A definition, benchmark and database of AI for social good initiatives. *Nat Mach Intell*, 3:111–115.

Daniel, F., Cappiello, C. & Benatallah, B. 2019. Bots acting like humans: Understanding and preventing harm. *IEEE Internet Computing*, 23(2):40-49.

De Jong, P. 2015. Solution-focused Therapy. In: Corcoran, K. & Roberts, A.R., eds. *Social Workers' Desk Reference*. 3rd ed. New York: Oxford University Press. pp. 268-275.

Denecke, K., Abd-Alrazaq, A., & Househ, M. 2021. Artificial Intelligence for Chatbots in Mental Health: Opportunities and Challenges. In: Househ, M., Borycki, E., Kushniruk, A. eds. *Multiple Perspectives on Artificial Intelligence in Healthcare*. Cham: Springer. pp. 115-128.

Dhyani, M. and Kumar, R., 2021. An intelligent Chatbot using deep learning with Bidirectional RNN and attention model. *Materials today*, 34(3):817-824.

El-Ansari, A. & Beni-Hssane, A. 2023. Sentiment Analysis for Personalized Chatbots in E-Commerce Applications. *Wireless Personal Communications*, 129(3):1623-1644.

Fang, L., Tarshis, S., McInroy, L. & Mishna, F. 2018. Undergraduate Student Experiences with Text-Based Online Counselling. *British Journal of Social Work*, 48(6):1774-1790.

Fava, N.M., Schuler, B.R. & Kelly, M.S. 2016. The incorporation of big data in social work practice: Current efforts and future directions. *Journal of Social Work Education*, 52(3):316-328.

Finlayson, B.T., Jones, E. and Pickens, J.C. 2023. Solution focused brief therapy telemental health suicide intervention. *Contemporary family therapy*, 45(1):49-60.

Foster, I., Ghani, R., Jarmin, R.S., Kreuter, F. & Lane, J. 2020. *Big Data and social science: Data science methods and tools for research and practice*. CRC Press.

Giglio, F. 2021. Fintech: A literature review. *European Research Studies Journal*, 24(2B):600-627.

Giguère-Rancourt, A., Plourde, M., Racine, E., Couture, M., Langlois, M., … Simard, M. 2022. Goal management training and psychoeducation / mindfulness for treatment of executive dysfunction in Parkinson's disease: A feasibility pilot trial. *PLoS ONE*. 17(2):e0263108.

Go, E. & Sundar, S.S. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97:304-316.

Goldkind, L. 2021. Social Work and Artificial Intelligence: Into the Matrix. *Social Work,* 66(4):372–374.

Gupta, A., Hathwar, D. & Vijayakumar, A., 2020. Introduction to AI chatbots. *International Journal of Engineering Research and Technology*, 9(7):255-258.

Grudin, J., & Jacques, R. 2019. Chatbots, humbots, and the quest for artificial general intelligence. *Conference proceedings*. 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland. New York: Association for Computing Machinery. pp. 1-11.

Güzel, B. 2021. Futurism in Social Work: Transhumanism, Gamification of Human Life, Homo Roboticus and Care. *Journal of Society & Social Work*, 32(2):589-606.

Hajjaji, Y., Boulila, W., Farah, I.R., Romdhani, I. & Hussain, A. 2021. Big data and IoT-based applications in smart environments: A systematic review. *Computer Science Review*, 39:1-39.

Hardeniya, N.H., Perkins, J., Chopra, D., Joshi, N. & Mathur, I. 2016. *Natural Language Process: Python and NLTK*. Birmingham: Pact.

Hassan, S.U., Ahamed, J. & Ahmad, K. 2022. Analytics of machine learning-based algorithms for text classification. *Sustainable operations and computers*, 3:238-248.

Heath, S., Liddle, J. & Wiles, J. 2019. The Challenges of Designing a Robot for a Satisfaction Survey: Surveying Humans Using a Social Robot*. International Journal of Social Robotics*, 12:519-533.

Hiland, E.B. 2021. *Therapy Tech: The Digital Transformation of Mental Healthcare*. Minneapolis, MN: University of Minnesota Press.

Holzmeyer, C., 2021. Beyond 'AI for Social Good'(AI4SG): social transformations—not tech-fixes—for health equity. *Interdisciplinary Science Reviews*, 46(1-2):94-125.

Hurbans, R. 2020. *Artificial Intelligence Algorithms*. New York, NY: Manning Publications.

Hussain, M. & Manhas, J. 2016. Artificial intelligence for big data: Potential and relevance. *International Academy of Engineering and Medical Research*, 1(1):1-5.

Inkster, B., Sarda, S. & Subramanian, V. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR Mhealth Uhealth*, 6(11):3-18.

Jafari, H., Kassan, A., Reay, G., & Climie, E. A. 2022. Resilience in Refugee Children and Youth: A Critical Literature Review. *Canadian Psychology/Psychologie canadienne,* 3, art. #09. doi: 10.31703/glr.2018(III-I).09

Jain, M., Kumar, P., Kota, R. & Patel, S.N. 2018. *Evaluating and informing the design of chatbots*. Paper delivered at the 2018 Designing interactive systems conference, China, Hong Kong. http://dx.doi.org/10.1145/3196709.3196735 Date of access: 16 Dec. 2022.

Javaid, M., Haleem, A., Singh, R.P. & Suman, R. 2022. Artificial intelligence applications for industry 4.0: A literature-based study. *Journal of Industrial Integration and Management*, 7(1):83-111.

Johnson, A. 2023. ChatGPT In Schools: Here's Where It's Banned—And How It Could Potentially Help Students. *Forbes*, 18 Jan. https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students/?sh=7c237b7e6e2c Date of access: 28 Jan. 2023.

Johnson, E.T., Huck, G., Brooks, J. Moser, E. Blake, J. & Chan, F. 2015. Solution-Focused Brief Therapy. In: Chan, F., Berven, N.L. & Thomas, K.R., eds. *Counseling Theories and Techniques for Rehabilitation and Mental Health Professionals*. New York: Springer. pp. 49-69.

Jones, C.L.E., Hancock, T., Kazandjian, B. & Voorhees, C.M. 2022. Engaging the Avatar: The effects of authenticity signals during chat-based service recoveries. *Journal of Business Research*, 144:703-716.

Joulin, A., Grave, E., Bojanowski, P. & Mikolov, T. 2016. Bag of tricks for efficient text classification. https://arxiv.org/abs/1607.01759 Date of access 31 Dec. 2022.

Jung, S. 2019. Semantic vector learning for natural language understanding. *Computer Speech & Language*, 56:130-145.

Keenan, M.J.S. 2019. *Advanced Positioning, Flow, and Sentiment Analysis in Commodity Markets : Bridging Fundamental and Technical Analysis*. Newark, NJ: John Wiley & Sons.

Khan, Z.F. & Alotaibi, R.A. 2020. Applications of Artificial Intelligence and Big Data Analytics in m-Health: A Healthcare System Perspective. *Journal of Healthcare Engineering*, 2020, art. #8894694. doi:10.1155/2020/8894694

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Bhale, P. & Das, T. 2015. A study of today's A.I. through chatbots and rediscovery of machine intelligence. *International Journal of u- and e- Service, Science and Technology*, 8(7):277-284.

Khurana, D., Koli, A., Khatter, K., & Singh, S. 2022. Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 37, art.#14. doi:10.1007/s11042-022-13428-4

Kumar, V.M., Keerthana, A., Valliammai, S. & Vinithasri, V. 2016. Sanative Chatbot For Health Seekers. *International Journal Of Engineering And Computer Science,* 5(3):16022-16025.

Kumar, A. & Kumar, A. 2022. Human Sentiment Analysis on Social Media through Naïve Bayes Classifier. *Journal of Scientific Research*, 66(1):350-357.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A. & Singh, I. 2019. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights*, 11:1-9.

Lee, C.T., Pan, L.Y. & Hsieh, S.H. 2022. Artificial intelligent chatbots as brand promoters: a two-stage structural equation modelling-artificial neural network approach. *Internet Research*, 32(4): 1329-1356.

Lipman, J. & Distler, R. 2023. Schools Shouldn't Ban Access to ChatGPT. *Time*, 10 Jan. https://time.com/6246574/schools-shouldnt-ban-access-to-chatgpt/ Date of access: 28 Jan. 2023.

Lommatzsch, A. & Katins, J. 2019. An Information Retrieval-based Approach for Building Intuitive Chatbots for Large Knowledge Bases. *Lernen, Wissen, Daten, Analysen*, 49(2):343-352.

Lucey, B. & Dowling, M. 2023. ChatGPT: our study shows AI can produce academic papers good enough for journals – just as some ban it. *The Conversation*, 26 Jan. https://theconversation.com/chatgpt-our-study-shows-ai-can-produce-academic-papers-good-enough-for-journals-just-as-some-ban-it-197762 Date of access: 28 Jan. 2023.

Luger, E. & Sellen, A. 2016. *Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents*. Paper delivered at the 2016 CHI Conference on Human Factors in Computing Systems, USA, New York. doi:10.1145/2858036.2858288

Marsch, L.A. 2021. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacol.* 46:191–196.

McCarthy, J. 1956. Measures of the value of information. *Proceedings of the National Academy of Sciences*, 42(9):654-655.

McCarthy, J., Minsky, M.L., Rochester, N., & Shannon, C.E. 1956. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.

McGee, V. 2022. *Virtual Therapy vs. In-Person Therapy.* https://www.psychology.org/resources/virtual-therapy-vs-in-person/ Date of access: 06 Apr. 2022.

Mehanna, Y.S. & Mahmuddin, M. 2021. The Effect of Pre-processing Techniques on the Accuracy of Sentiment Analysis Using Bag-of-Concepts Text Representation. *SN Computer Science*, 2:237-250.

Mehta, P. & Pandya, S. 2020. A review on sentiment analysis methodologies, practices and applications. *International Journal of Scientific and Technology Research*, 9(2):601-609.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. & Dean, J. 2013. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26:1-8.

Mills, K.A. 2019. *Big Data for Qualitative Research*. New York, NY: Routledge.

Mierzwa, S., Souidi, S., Conroy, T., Abusyed, M., Watarai, H., & Allen, T. 2019. On the Potential, Feasibility, and Effectiveness of Chat Bots in Public Health Research Going Forward. *Online journal of public health informatics*, 11(2), e4. doi:10.5210/ojphi.v11i2.9998

Mindell, D.A. & Reynolds, E. 2023. *The work of the future: building better jobs in an age of intelligent machines*. Cambridge, MA: MIT Press.

Molano, J.I.R., Lovelle, J.M.C., Montenegro, C.E., Granados, J.J.R. & Crespo, R.G. 2018. Metamodel for integration of internet of things, social networks, the cloud and industry 4.0. *Journal of ambient intelligence and humanized computing*, 9:709-723.

Molnar, A. & de Shazer, S. 1987. Solution-focused therapy: Toward the identification of therapeutic tasks. *Journal of Marital and Family Therapy*, 13(4):349-358.

Moor, J. 2006. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. AI Magazine, 27:87-91.

Morrissey, D. 2021. *Social Work and Helping Humanity Through Chatbots*. https://medium.com/@devinmorrissey.1987/social-work-and-helping-humanity-through-chatbots-a669dab63b4d Date of access: 14 Jan. 2023.

Nandwani, P. & Verma, R. 2021. A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1):81-100.

Nolan, B. 2023. A freelance writer says ChatGPT wrote a R10,000 article in just 30 seconds. *Business Insider South Africa*, 25 Jan. https://www.businessinsider.co.za/chatgpt-wrote-600-dollar-article-30-seconds-freelance-writer-ai-2023-1 Date of access: 28 Jan. 2023.

OpenAI. 2023. Join us in shaping the future of technology. https://openai.com/ Date of access: 28 Jan. 2023.

Patel, K.K., Patel, S.M. & Scholar, P. 2016. Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. *International journal of engineering science and computing*, 6(5):6122-6131.

Pencheva, I., Esteve, M. and Mikhaylov, S.J. 2020. Big Data and AI–A transformational shift for government: So, what next for research?. *Public Policy and Administration*, 35(1):24-44.

Porter, N.D., Verdery, A.M. & Gaddis, S.M. 2020. Enhancing big data in the social sciences with crowdsourcing: Data augmentation practices, techniques, and opportunities. *PLoS ONE*, 15(6), e0233154. https://doi.org/10.1371/journal. pone.0233154

Pradhan, A., Senapati, M. R. & Sahu, P. K. 2022. Improving Sentiment Analysis with Learning Concepts from Concept, Patterns Lexicons and Negations. *Ain Shams Engineering Journal*, 13(2), art. #101559. doi:10.1016/j.asej.2021.08.004

Prochaska, J.J., Vogel, E.A., Chieng, A., Kendra, M., Baiocchi, M., Pajarito, S. & Robinson, A. 2021. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. *Journal of medical Internet research*, 23(3), e24850. doi:10.2196/24850

Rice, E., Wilder, B., Onasch-Vera, L., DiGuiseppi, G., Petering, R., Hill, C., Yadav, A., Lee, S.J. & Tambe, M. 2021. A Peer-Led, Artificial Intelligence–Augmented Social Network Intervention to Prevent HIV Among Youth Experiencing Homelessness. *Journal of Acquired Immune Deficiency Syndromes*, 88(1):20-26.

Rivas, P., Holzmayer, K., Hernandez, C. & Grippaldi, C. 2018. Excitement and concerns about machine learning-based chatbots and talkbots: A survey. *Conference proceedings*. IEEE International Symposium on Technology and Society (ISTAS), Washinton, DC, United States of America. Institute of Electrical and Electronics Engineer. pp. 156-162.

Roestenburg, W.J.H. 2021. *Analysing technology acceptance by social workers in South Africa*. Potchefstroom: North-West University. (Dissertation – M.B.A.).

Russel, S.J. & Norvig, P. 2022. *Artificial Intelligence: A Modern Approach*. 4th ed. Harlow: Pearson Education.

Safodien, M. 2021. Social work 4.0? The fourth industrial revolution and social work education: a South African perspective. *Social Work/Maatskaplike Werk*, 57(3):257-271.

Sage, S. & Singer, J.B. 2022. Technology and Social Work Practice. Micro, Mezzo and Macro Applications. In: Rapp-McCall, L., Roberts, A. & Corcoran, K., eds. *Social Workers' Desk Reference*. New York, NY: Oxford University Press. pp. 45-52.

Sánchez, P.M., Aburto, L.G., Orellana, M.O. & Torres, R.C. 2020. Knowledge and skills for occupational social work: challenges for education and professional practice. *Cuadernos de Trabajo Social*, 33(1):247-258.

Sankar, R. 2018. Empowering chatbots with business intelligence by big data integration. *International Journal of Advanced Research in Computer Science*, 9(1):627-631.

Santos, G.A., De Andrade, G.G., Silva, G.R.S., Duarte, F.C.M., Da Costa, J.P.J. & De Sousa, R.T. 2022. A Conversation-Driven Approach for Chatbot Management. *IEEE Access, 10:8474-8486. doi:*10.1109/ACCESS.2022.3143323

Santos, M.Y. & Costa, C. 2020. *Big Data: Concepts, Warehousing, and Analytics*. New York, NY: Routledge.

Sauer, P.C. & Seuring, S. 2023. How to conduct systematic literature reviews in management research: a guide in 6 steps and 14 decisions. *Review of Managerial Science*, 17(5):1899-1933.

Setyawan, M.Y.H., Awangga, R.M. & Efendi, S.R. 2018. Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot. *Conference proceedings*. 2018 International Conference on Applied Engineering (ICAE, 2018), Batam, Indonesia. Institute of Electrical and Electronics Engineers. pp. 1-5.

Shum, H.Y., He, X.D. & Li, D. 2018. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*:10-26.

Shreyas, J., Jumnal, A., Kumar, S.D. & Venugopal, K.R. 2020. Application of computational intelligence techniques for internet of things: an extensive survey. *International Journal of Computational Intelligence Studies*, 9(3):234-288.

Sinha, S.K. & Dubey, A.D. 2022. Artificial Intelligence Concerning Corporate Sustainability: A Systematic Review of Research. In: Goel, R., Baral, S.K. & Venkatesh, R., eds. Internet of Things and Machine Learning. New York, NY: Nova Science Publishers. pp. 215-233.

Smys, S. & Haoxiang, W. 2021. Naïve Bayes and entropy based analysis and classification of humans and chat bots. *Journal of IoT in Social Mobile Analytics and Cloud*, 3(1):40-49.

Solanki, P., Grundy, J. & Hussain, W. 2023. Operationalising ethics in artificial intelligence for healthcare: A framework for AI developers. *AI and Ethics*, 3(1):223-240.

Srinivasan, J. 2021. *2 Ways AI is Transforming Wellness in the Workplace*. https://www.spiceworks.com/hr/engagement-retention/guest-article/2-ways-ai-is-transforming-wellness-in-the-workplace/ Date of access: 14 Jan. 2023.

Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J… 2019. *Mitigating gender bias in* natural language processing: Literature review. https://arxiv.org/abs/1906.08976 Date of access: 30 Mar. 2023.

Sun, T.X., Liu, X.Y., Qiu, X.P. & Huang, X.J. 2022. Paradigm shift in natural language processing. *Machine Intelligence Research*, 19(3):169-183.

Sunter, C. & Ilbury, M. 2021. *Thinking the future: New Perspectives from the shoulders of giants.* Cape Town: Penguin Books.

Suta, P., Lan, X., Wu, B., Mongkolnam, P. & Chan, J.H. 2020. An overview of machine learning in chatbots. *International Journal of Mechanical Engineering and Robotics Research*, 9(4):502-510.

Taecharungroj, V. 2023. “What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1):35.

Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T. and Qin, B. 2014. Learning sentiment-specific word embedding for twitter sentiment classification. *Conference proceedings.* 52nd Annual Meeting of the Association for Computational Linguistics (ACL, 2014), Baltimore, United States of America. Baltimore: Association for Computational Linguistics. pp. 1555-1565.

Tambe, M. & Rice, E. 2018. Merging Social Work Science and Computer Science for Social Good. In: Tambe, M & Rice, E., eds. *Artificial Intelligence and Social Work*. Cambridge: Sheridan Books. pp. 48-72.

Tebenkov, E. and Prokhorov, I. 2021. Machine learning algorithms for teaching AI chatbots. *Procedia Computer Science*, 190:735-744.

Terblanche, L. 2021. Employee assistance programmes and occupational social work: interrelated with distinct features. *Social Work/Maatskaplike Werk*, 57(1):16-38.

Turing, A.M. 1950. Computing Machinery and Intelligence. *Mind,* 59(236):433–60.

Vaidyam, A.N., Wisniewski, H., Halamka, J.D., Kashavan, M.S. & Torous, J.B. 2019. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7):456-464.

Valério, A.M., Guimarães, T.G., Prates, R.O. & Candello, H. 2017. *Here's What I Can Do: Chatbots' Strategies to Convey Their Features to Users.* Paper delivered at the XVI Brazilian Symposium on Human Factors in Computing Systems, Joinville. https://dl.acm.org/doi/10.1145/3160504.3160544 Date of access: 18 Dec. 2022.

Vijayaraghavan, V., Cooper, J.B. & Leevinson, R.J. 2020. Algorithm inspection for chatbot performance evaluation. *Procedia Computer Science*, 171:2267-2274.

Vilaza, G.N. & McCashin, D. 2021. Is the Automation of Digital Mental Health Ethical? Applying an Ethical Framework to Chatbots for Cognitive Behaviour Therapy. *Frontiers in Digital Health*, 3, art. #2673. doi:10.3389/fdgth.2021.689736

Visvikis, D., Le Rest, C.C., Jaouen, V. & Hatt, M. 2019. Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications. *European Journal of Nuclear Medicine and Molecular Imaging*, 46:2630-2637.

Wallace, R.S. 2009. The anatomy of ALICE. In: Epstein, R. & Peters, G., eds. *Parsing the Turing test*. Dordrecht: Springer. pp. 181-210.

Weizenbaum, J. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36-45.

Wu, E.H.K., Lin, C.H., Ou, Y.Y., Liu, C.Z., Wang, W.K. & Chao, C.Y. 2020. Advantages and constraints of a hybrid model K-12 E-Learning assistant chatbot. *IEEE Access,* 8:77788-77801. doi:10.1109/ACCESS.2020.2988252

Wulf, A. 2020. How to use text messages as a dialogue tool in Solution Focused conversations. In: Dierolf, K., Hogan, D. & van der Hoorn, S., eds. *Solution Focused Practice Around the World*. New York: Routledge. pp. 94-107.

Yang, H., Lee, Y.I., Lee, H.J., Cho, S.W. & Koo, M.W. 2015. A study on word vector models for representing Korean semantic information. *Phonetics and Speech Sciences,* 7(4):41-47.

Yarowsky, D. 1995. Unsupervised word sense disambiguation rivalling supervised methods. Paper delivered at the 33rd annual meeting of the association for computational linguistics, Cambridge, MA. https://aclanthology.org/P95-1000/ Date of access: 14 Jan. 2023.

Yates, T.M. & Frank, L.S. 2021. Natural language processing: Opportunities and challenges for social work research. *Social Work Research*, 45(3):145-156.

Zetino, J. & Mendoza, N. 2019. Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare. *Social Work in Public Health*, 34(5):409-417.

Zhu, Y., Janssen, M., Wang, R. & Liu, Y. 2022. It is me, chatbot: working to address the COVID-19 outbreak-related mental health issues in China. User experience, satisfaction, and influencing factors. *International Journal of Human–Computer Interaction*, 38(12):1182-1194.

Zumstein, D. & Hundertmark, S. 2017. Chatbots - An Interactive Technology for Personalized Communication, Transactions and Services. *IADIS International journal on www/internet*, 15(1):96-109.

CHAPTER 3: EMPIRICAL FINDINGS

3.1. INTRODUCTION

The 4IR is a reality that social work practitioners, particularly those involved in EWHPs, must address to ensure continued relevance. The field's relevance will depend on how practitioners, educators, policymakers, researchers, and organisational leaders integrate AI advances, focusing on policy, ethics, and the need for ongoing professional development to keep pace with technological advancements (Mathekga & Sekudu, 2019). Ongoing professional development involves equipping social work practitioners with new skills, knowledge, and competencies required to effectively utilise AI tools while adapting to the evolving technological landscape. These stakeholders must thoughtfully incorporate AI-driven systems that support ethical decision-making, maintain professional standards, and uphold social justice while offering adaptable services. By fostering human-AI collaboration, these stakeholders can ensure that social work remains engaged with emerging technologies, whilst addressing the demands and challenges of the 4IR (Macnab-Holding, 2024; Ndzendze et al., 2024).

Researchers must be mindful of the unavoidable obstacles arising from the 4IR. Thus, this study aimed to introduce and construct a novel conceptual framework for an AI-driven generative chatbot counselling system. The framework's primary objective is to enhance a multi-faceted AI algorithm's capabilities within the domain of social work. This research specifically centred on catering to the requirements of social workers engaging with a comprehensive EHWP.

3.2. RATIONALE FOR STUDY

The IoT is a pivotal element driving the 4IR, bridging the gap between physical and digital applications. At its core, the IoT refers to the interconnectivity of commodities and people facilitated by interconnected technology and diverse platforms (Giglio, 2021; Hajjaji *et al.*, 2021). Ashton (1999) has introduced fundamental concepts of the IoT, expanding device-to-device connectivity on a larger scale than previously achievable. Semantically, Ashton (2009:97) defines the IoT as a "global network of interconnected objects". Building upon this, Patel *et al.* (2016) elaborate that the IoT encompasses a dynamic global network infrastructure, utilising Internet connectivity and specific agreement protocols to identify, control, and monitor systems. This is achieved through the integration of physical and virtual objects enabled by interoperability. Common examples of the IoT in everyday life include wearable devices, self-driving cars, crop monitors, logistic tracking, and retail inventory tracking (Shreyas *et al.*, 2020).

The IoT revolves around enhancing the provision and reception of services for customers, individuals, and businesses. Technological advancements have prompted organisations to adopt novel approaches to managing and valuing their assets. The convergence of AI and IoT, as

highlighted by Javaid *et al.* (2022), holds immense potential to generate impactful applications that enhance various aspects of daily life. Further innovative and groundbreaking applications are anticipated as AI and IoT continue to progress. The incorporation of AI is integral to new operational models, compelling businesses and organisations to reassess their operations and strive for increased agility and effectiveness. As in other scientific domains, social work and EHWPs have experienced the impact of technological advancements, with technology providing diverse forms of support for mental health (Shreyas *et al.*, 2020).

While the integration of big data technologies into social work is yet to be implemented in South Africa, it is actively being pursued in foreign contexts, particularly in areas such as identifying patterns of homelessness, healthcare and social services planning, risk assessment through predictive modelling, program evaluation, and evidence-based practice (Foster *et al.*, 2020). Future applications of big data-driven social work are expected to capture, update, and respond to real-time data on human behaviour and needs (Fava *et al.*, 2016; Marsch, 2021).

AI and social work may initially appear to be unrelated fields; however, Tambe and Rice (2018) argue that their fusion can create a dynamic new scientific domain. The exclusive collaborations inherent to both fields contribute to the emergence of AI4SG or AI and social work as a distinct intellectual arena. Given the current advancements in AI and DL, chatbots offer a promising augmentation to traditional service delivery.

This study proposed the integration of chatbot interactions into traditional practice, rather than advocating for the replacement of AI technology and mental health professionals. However, the field of social work currently lacks AI-specific therapy platforms designed for its unique requirements in the context of the 4IR, automation, AI, and IoT. Social work needs to adapt and expand its problem-solving methodology in response to the impending ubiquity and cost-effectiveness of digital-led technologies (Sunter & Ilbury, 2021).

Chatbots present a novel and interactive modality for engaging with individuals seeking mental health services and may serve as a potential solution. This study, therefore, aimed to enrich the knowledge base and theoretical framework of social work. Successful implementation could contribute to the modernisation of the field, enabling its relevance in the face of the 4IR, automation, and IoT. The ultimate objective of this project was to create an extensive system of platforms and tools powered by AI that integrate naturally into social work practice.

3.3. CURRENT USE OF AI IN SOCIAL WORK AND EHWP

The application of AI in the domains of social work and EHWPs is an emerging field with significant potential. A notable use of AI technology is in the development of chatbots. Safodien (2021) underscores the importance of social work evolving to remain pertinent in the era of the 4IR. Güzel

(2021) also places a strong emphasis on the integration of AI and technology into social work practices and educational curricula.

The concept of AI4SG seeks to leverage AI to address societal challenges and enhance overall wellbeing. Research on AI approaches has shown promise in addressing societal issues, such as raising awareness about HIV prevention among homeless youth (Rice *et al.*, 2021). Major corporations are increasingly investing in this endeavour as they recognise the potential of AI4SG. Scholars, governments, corporations, and non-profit organisations should, therefore, collaborate in the advancement of AI4SG.

Goldkind (2021) states that social work must engage with algorithmic and AI entities to expand the definitions of 'individual' and 'environment'. While AI solutions are not a universal remedy, they can serve as valuable tools for addressing critical social concerns (Cowls *et al.*, 2021).

According to Roestenburg (2021), social workers are receptive to the utilisation of IT technology provided that it aligns with market demands. It is advisable to understand these market demands, adapt to the constantly evolving information landscape, and acknowledge the ongoing nature of technology adoption.

Morrissey (2021) draws attention to the unregulated use of AI algorithms, particularly in the assessment of child welfare and the implementation of chatbots in social work. Furthermore, Inkster *et al.* (2018) have studied the impact of a counselling chatbot (Wysa) on the emotional wellbeing of workers, suggesting that it can be a valuable resource in situations involving EHWPs.

There is an expectation that AI will have a profound influence on both professionals working in the psychosocial field and organisations dedicated to wellness. Baldwin (2019) projects that within approximately 50 years, AI will be able to perform critical social skills at a level comparable to humans. According to Srinivasan (2021), AI-driven chatbots are already revolutionising how workers access mental health resources, particularly among younger generations, where chatbots are initiating conversations.

The domain of chatbots in social work and EHWPs is in its early stages but holds significant promise. Research in the AI4SG field is expanding, with a primary focus on harnessing AI to address societal issues. Thus, the role of AI in social work is evolving, with collaboration and technological adaptation being essential components.

3.4. METHODOLOGY AND RESEARCH DESIGN

A distinct approach to social work research methodology was formulated to fully execute the essential chatbot design methodologies encompassing text analysis, data pre-processing, input embedding, NLP, the Naïve Bayes classifier, ML, sentiment analysis and ANNs. To accurately and methodically define the characteristics of the phenomenon under investigation, the study

employed a descriptive design. A deeper understanding of the data was made possible by recognising, assessing, and synthesising patterns and themes that emerged during the investigation. The mixed-method approach was improved by the descriptive design, which made it easier to clearly illustrate qualitative themes and, when appropriate, permitted the addition of measurable aspects (Doyle *et al*., 2019). The design consisted of a multi-phased, mixed-method approach in the form of a qualitative first, quantitative second data analysis, incorporating thematic analysis (qualitative analysis – section 3.5.), sentiment analysis (quantitative analysis – section 3.6.), and word frequency analysis (quantitative analysis – section 3.7.).

Live chat transcripts for six months, consisting of 50 or more interactions between human counsellors and clients of Lyra SA that took place on the platform Cubit, were subjected to thematic and textual analysis. Thematic analysis was employed to identify underlying themes and patterns within the chat transcripts, while word frequency analysis provided insights into the most commonly used words and their significance. Additionally, sentiment analysis was conducted to gauge the emotional tone and sentiment expressed in the interactions. This comprehensive design has resulted in the acquisition of data that is essential for informing the aforementioned components of an intelligent online counselling framework for an EHWP. It is important to note that these excerpts have been represented exactly as they were transcribed, with certain cadences and spelling errors included. This was done to inform the reader of the tense and tone that users and counsellors made use of in these conversations.

3.5. EMPIRICAL DATA: LIVE CHAT TRANSCRIPT ANALYSIS

This section examined the empirical data of the live chat transcript analysis. By analysing live chat session transcripts, patterns emerged that illustrate the flow and structure of sessions, the concerns that users presented, and the therapeutic approaches followed by live chat counsellors. Each theme represents a distinct aspect of these engagements. The following five themes are examined in the subsequent section: i) flow/process flow of live chat sessions; ii) problems/challenges; iii) counsellor approaches; iv) linguistic trends; and v) technical issues.

3.5.1. Theme 1: Flow/process flow of live chat sessions

The live chat transcripts that were analysed (N=90) followed a similar pattern in terms of the flow of the therapy sessions. The flow of the counselling process consisted of numerous textual exchanges between the agent and the user. A typical process flow would involve the client initiating the session employing some form of greeting, which could either be undetailed and/or concise: *[2(136) - User: hi | Agent: Good morning | User: i need assistance working through my childhood trauma and abuse. So that I do not keep repeating bad patterns with my own child]* or comprehensive and detailed:

[2(250) – User: Hi, I am in a very difficult situation financially, i am over indebted and applied for debt review. The process has been tedious. My main issue is that I am battling with anxiety and depression and sleepless nights because I cannot accept how I have gotten myself here. I am married with four kids, I live with my parents, I have siblings and in-laws that are unemployed and helped them for many years, basically I was the Pillar, the good one, the one who wanted to please everyone and could not say no, ive created this character of someone who they can always turn to help and made loans after loans and credit cards without anyones knowledge, not even my husband. In December all this debt caught up to me and I could no longer manage it, i went into deep states of anxiety attacks and depression where I am sure my family noticed. I stopped doing many things I usually would do and I canten hide it. My husband now knows the truth about my debts and has been supportive in allowing me to fix this by debt review. Thing is my parents who depend on me does not know and I am too afraid to tell them because they think so highly of me and always know they can turn to me for anything. I have siblings who are also struggling and I have been supporting them too for many years. But i have done it all wrong for years to a point that I now have to be in debt review. I enjoyed always having to help everyone too much and I have nothing to even show for all the debts I have, I went on sprees just providing, paying fees, bdays, food in the house, I have not been smart to secure long term assets like a home or car, nothing, just LIVING things for daily needs and extras wanting my kids to have everything they need. I cannot forgive myself since ive come to the realization of losing my salary to debts and not having back up to make provision. Resulting in the anxiety and panick attacks, insomnia and depression. I dont know when last I had 8 hours of sleep. How do I deal with this?].

Once the agent received the greeting, the agent could respond to the client*: [2(384) – User: Hi Agent: Hello, how may I assist you today?].* The analysis found that the agent, on occasion, would utilise a built-in functionality that sends an automated greeting:

[2(513) – Agent: Thank you for making contact with Lyra SA. Before we begin, I would like to introduce myself and explain the process to you. I am a counsellor. I will gladly assist you. All our communication is completely secure and confidential on this chat platform. As a first step, I would like to find out what your best hopes are for our chat today?].

If the user did not have a case with Lyra SA, the agent could request the user's biographical information*: [2(87) – Agent: Could you kindly provide me with the following details (in order to create a reference number for yourself); Your full name; The name of your Company; Company Division; Job title; Employee Number], which are needed for statistical and case management purposes].*

After the user provided their biographical information, the agent typically commenced to probing the reason for the user's engagement: *[2(294) – Agent: Have you spoken to your husband regarding his problems].* The user typically responded with either a concise or comprehensive answer. For example, *[2(294) – User: yes, after which the therapeutic session transitions into a complete session]* or:

[2(294) – User: he needs help. he has an alcohol problem. feel so stressed out because my mother on the other side keeps mentioning the same stories over and over just feel like it's becoming too much. feel like I'm getting anxiety and feel so depressed . he doesn't acknowledge that he has an alcoholic.problem. he is a different person when he is sober we hardly even talk].

Once the situation at hand had been probed, the agent would draw upon therapeutic techniques, which usually involved goal setting and desired outcomes of the session: *[(2)117 - Agent: The first step is to set up some goals, expectations for your counselling. This means what you would want to gain from the counselling perspective].* At this point, the therapist engaged in the full use of their therapeutic knowledge and skills and the therapy session ran its course. During this stage, the therapist actively applied their extensive therapeutic expertise and aptitude, while allowing the therapy session to progress naturally. This process is illustrated in Figure 14.

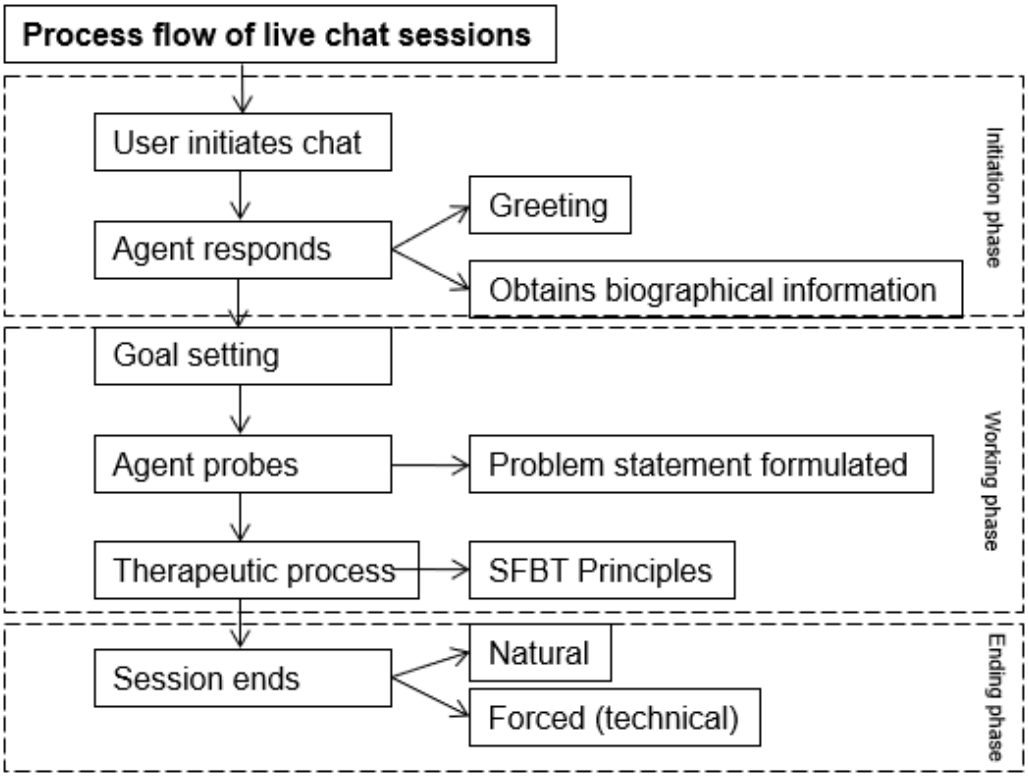


Figure 14. Process flow of live chat sessions

3.5.2. Theme 2: Problems/challenges

Users connected to seek assistance for a broad spectrum of problems or difficulties, which often involved difficulties in their personal and professional lives. As shown in Table 5, problems can be categorised as career and work-related concerns, dysfunction, emotional wellbeing, family dynamics, financial issues, interpersonal relationships, personal development, psychological challenges, and requests for miscellaneous help and support. The problem specifications (>4

occurrences) of the most frequent codes, which are included in the code groups in Table 5, are elaborately presented in Table 6 and Figure 16.

Table 5. Problem/challenge groupings

Code group	Count of code group
Career and Work-related Issues	174
Dysfunction	274
Emotional wellbeing	196
Family dynamics	101
Financial Issues	43
Interpersonal Relationships	96
Personal development	305
Psychological challenges	228
Request for help/support	193
Grand Total	**1610**

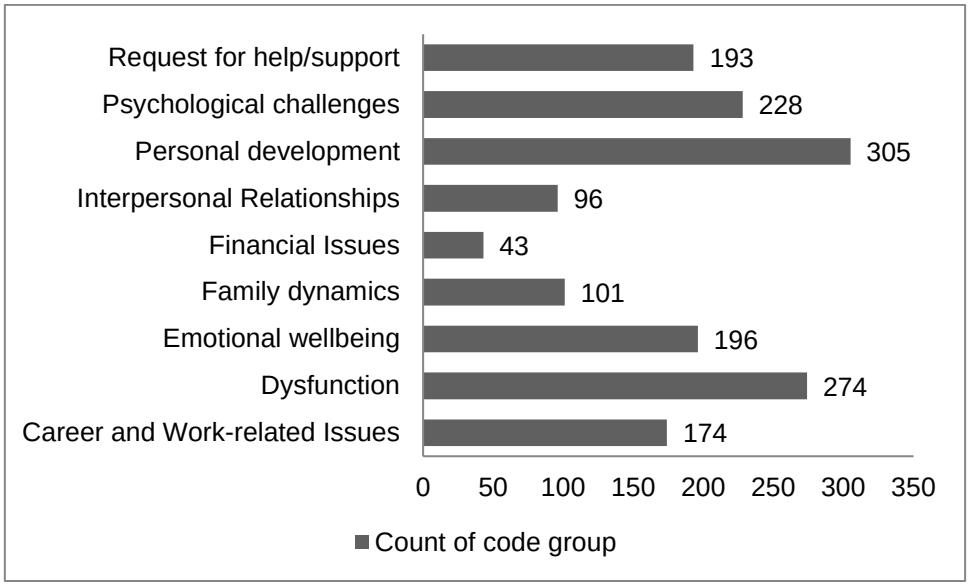


Figure 15. Count of code groups

Table 6 indicates the count of codes of co-occurring challenges or problems elaborately. These codes often co-occurred in the same live chat session and often appeared as or were similar to, different live chat sessions. Each of the most prominent codes (Figure 16) were assigned the value of sub-themes and are discussed based on their occurrences and merits.

Table 6. Occurrences of problems/challenges

Challenge/Problem	Count of codes
Personal development: General	305

Dysfunction	274
Psychological challenges: Uncategorised	228
Emotional wellbeing	196
Career and work-related: general	174
Interpersonal Relationships	104
Family dynamics	101
Personal development: Encouragement	59
Personal development: Empathy	39
Dysfunction: Uncertainty	22
Stress	19
Separation/divorce	17
Personal development: Gratitude	14
Personal development: Self-Care	13
Personal development: Encouragement, Personal development: Positive feedback	12
Fatigue: Uncategorised	11
Feeling overwhelmed	11
Interpersonal conflict: partner/spouse	11
Not coping	11
Psychological disorder: Suspected	11
Dysfunction: Uncertainty, Dysfunction: Questioning	10
Request for help/support: Suggestion	10
Experience of unpleasant emotions	9
Loss/bereavement	9
Anger	8
Emotional fatigue	8
Personal development: Self-improvement	8
Request for help/support: Advice	8
Childcare	7
Personal development: Reassurance	7
Interpersonal communication	7
Suicidal ideation	7
Religion	6
Trauma: Various	6
Dysfunction: Uncertainty, Dysfunction	6
Emotional wellbeing: Anxiety	6
Family dynamics: Family	6
Personal development: Positive reinforcement, Personal development: Encouragement	6
Abuse	5
Compassion fatigue	5
Dysfunction: Frustration	5
Dysfunction: Questioning	5
Emotional regulation	5

Emotional wellbeing: Anxiety, Dysfunction: Uncertainty	5
Family dynamics: Family relationships	5
Infidelity	5
No support system	5
Parental guidance	5
Career and Work-related Issues: Time management	4
Domestic violence	4
Emotional wellbeing: Mental health	4
Feeling “down”	4
Financial Issues: Financial stress, Family dynamics: Family issues	4
Personal development: Empathy, Request for help/support: Supportiveness	4
Personal development: Encouragement, Personal development: Motivation	4
Personal development: Encouragement, Personal development: Self-care	4
Personal development: Encouragement, Request for help/support: Suggestion	4
Personal development: Motivation	4
Personal development: Resilience	4
Physical abuse	4
Verbal abuse	4
Request for help/support: Problem-Solving	4
Emotional abuse	3
Total	**1875**

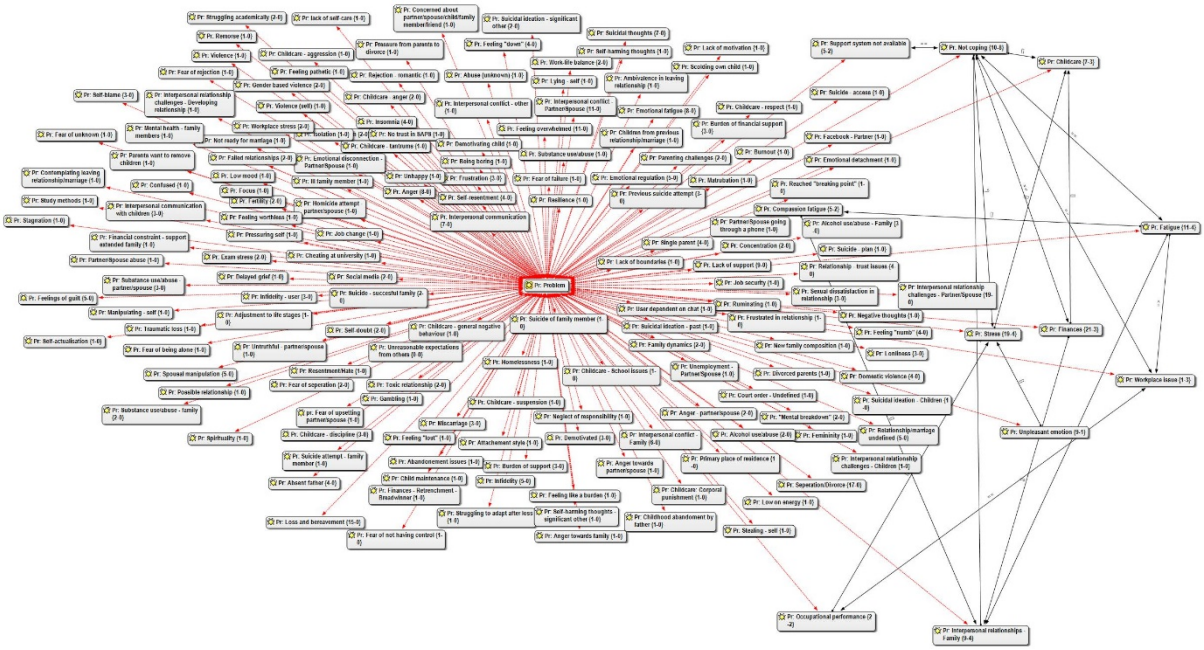


Figure 16. Problem network

3.5.2.1. Sub-theme 2.1. Personal development: General

Sub-theme 2.1 (Personal development: General) refers to a category of problems and challenges that users sought assistance for, relating directly or indirectly to their personal growth and self-improvement. In total, 305 codes occurred within this sub-theme. Codes that co-occurred include Boundaries {1-1}; Anger {8-2}; Anger - partner/spouse {2-4}; Anger towards partner/spouse {1-6}; Attachment style {1-2}; Being boring {1-2}; Cheating at university {1-1}; Childcare {7-5}; Concentration {2-3}; Confused {1-1}; Contemplating leaving relationship/marriage {1-4}; Demotivated {3-4}; Despondency {1-3}; Emotional detachment {1-3}; Emotional disconnection - Partner/Spouse {1-4}; Fear of being alone {1-2}; Fear of failure {1-1}; Fear of not having control {1-1}; Fear of rejection {1-6}; Fear of separation {2-3}; Fear of unknown {1-2}; Feeling pathetic {1-3}; Feeling worthless {1-3}; Feelings of guilt {5-3}; Focus {1-2}; Gambling {1-2}; Interpersonal communication {7-2}; Interpersonal communication with children {3-2}; Interpersonal conflict - family {6-2}; Interpersonal conflict - partner/spouse {11-5}; Interpersonal relationship challenges - developing relationship {1-2}; Isolation {1-2}; Lack of boundaries {1-1}; Lack of motivation {1-1}; Lack of self-care {1-1}; Loneliness {3-2}; Low on energy {1-7}; Lying - self {1-2}; Manipulating - self {1-2}; Negative thoughts {1-3}; New family composition {1-2}; Possible relationship {1-3}; Pressure from parents to divorce {1-1}; Pressuring self {1-1}; Rejection - romantic {1-2}; Relationship - trust issues {4-2}; Relationship/marriage undefined {5-36}; Self-resentment {4-4}; Study methods {1-2}; Toxic relationship {2-7}; Unpleasant emotion {9-3}, and Unreasonable expectations from others {0-1}. Excerpt 1 illustrates instances where reference has been made to Sub-theme 2.1 (namely, Personal development: General).

(2)19 - User: I just feel like I should be prioritizing myself, but I love him so much that its like I am being in denial about what we going through in order just be happy with him
(2)44 – User: it may not be perfect but I sacrificed so much for us till now I've given so much of myself I don' t know who I am anymore
2(87) – Use: The ideal outcome would be to conceive a child, and carry the pregnancy to full term with a live birth.
2(95) – User: …and there was a day where I she mixed up definitions and I assisted. and a day later she said to me that I must not use words I don't know cause ill look stupid
2(100) – User: …in order have my peace and be able to think better and remove myself from this situation because it's getting worse

Excerpt 1. Personal development: General

According to Lee *et al.* (2021), the wellbeing of individuals and society will be enhanced by ethical and knowledgeable AI. Self-awareness, emotional intelligence, resilience, self-improvement, and personal growth constitute parts of personal development, which encompasses an extensive spectrum of life experiences. It represents the search for personal improvement and a desire to

effect positive transformation (Bauer, 2021; Toyama *et al.*, 2020). Users requesting assistance under this sub-theme may have done so because they are experiencing the need to overcome personal challenges (which are too vast to list in this chapter), increase confidence, adopt healthier habits, improve self-esteem or focus on their general wellbeing, as displayed in Figure 17.

It is evident from the analysed live chat transcripts that users utilised live chat sessions to address a wide range of general personal development concerns, seeking guidance from live chat counsellors. The live chat counsellors, in turn, utilised a variety of therapeutic techniques (which will be discussed later in the chapter) to explore desired outcomes.

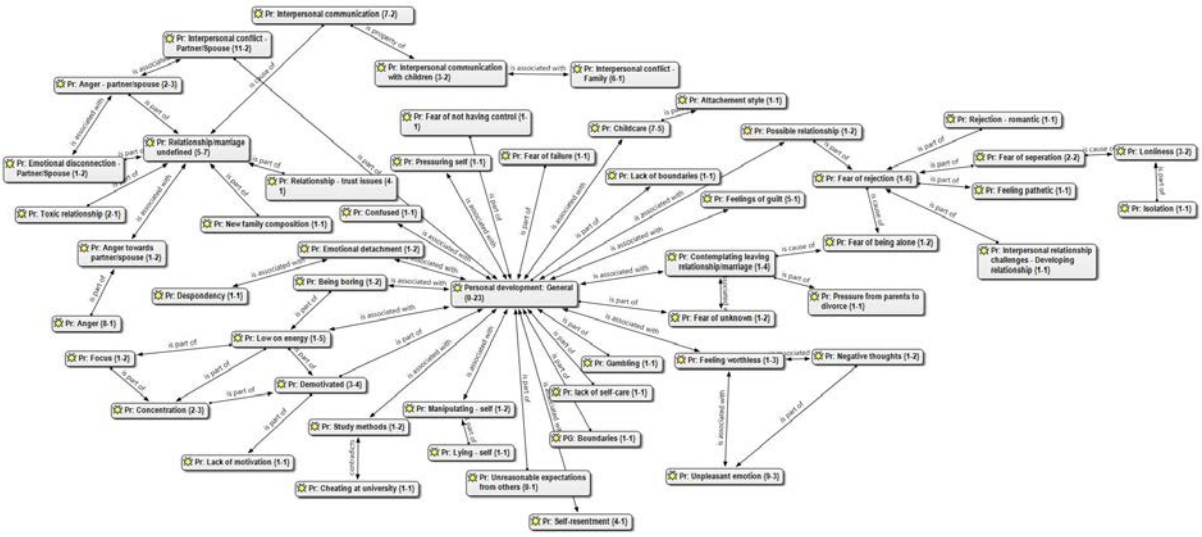


Figure 17. Sub-theme 2.1. Personal development: General network

3.5.2.2. Sub-theme 2.2. Dysfunction

In total, 274 codes occurred in Sub-theme 2.2 (Dysfunction). Codes that co-occurred include Emotional abuse {3-4}; Physical abuse {4-2}; Verbal abuse {4-3}; Addiction {0-1}; Aggression {0-1}; Criminal activities {0-1}; Divorce/Separation {0-2}; Suicide {0-12}; Dysfunction {0-10}; Alcohol use/abuse {2-1}; Alcohol use/abuse - Family {3-1}; Domestic violence {4-5}; Gambling {1-2}; Gender-based violence {2-2}; Homelessness {1-1}; Homicide attempt - partner/spouse {1-2}; Infidelity {5-6}; Previous suicide attempt {3-2}; Substance use/abuse {1-1}; Substance use/abuse - family {2-5}; Substance use/abuse - partner/spouse {3-2}; Suicidal ideation - Children {1-3}; Suicidal ideation - past {1-5}; Suicidal ideation - significant other {2-1}; Suicidal thoughts {7-7}; Suicide - access {1-4}; Suicide - plan {1-4}; Suicide - successful family {2-4}; Suicide attempt - family member {1-4}; Suicide of family member {1-3}, and Toxic relationship {2-7}. Excerpt 2 illustrates instances of interactions relating to Sub-theme 2.2 (Dysfunction).

2(294) – User: he needs help. he has an alcohol problem. feel so stressed out because my mother on the other side keeps mentioning the same stories over and over

2(229) – User: It's get my brother readmitted before he hurt either himself or one of us. I made a couple of attempts to have him admitted, I call the ambulance and the police for assistance from 12 o clock in the midnight when he had an episode trying to break through my door and shouting and screening, but could not get help.

2(236) – User: so 2 years ago I had the same problem and my dad hit me. my boyfriend at the time picked me up and I stayed with him and his family for about a year. last year i came back home . I felt alot more at ease when I was away. my only concern was my mum

2(250) – User: spot on, I am tired of being the good girl and provider (i am so tired of my name being the dialing name, the call out name) i want someone else to take the responsibility

2(262) – User: I now know - confirmed - she has a drug problem - at the time - it was that she needed legal help with a issue relating to her finances and that she would rectify once sorted. also financial help for living costs, gambling, cancer treatments

Excerpt 2. Dysfunction

Dysfunction in the context of Sub-theme 2.2. is defined by Kupferberg *et al.* (2016) as the impairment of social functioning, which is described as a person's capacity to carry out and fulfil typical social roles. It refers to a variety of issues that revolve around dysfunctional aspects within live chat user's personal and family lives, which manifest in various ways. The analysed transcripts, as displayed in Figure 18, indicated that the most common dysfunctionalities that users sought assistance for revolve around addiction issues, family conflict, abuse, aggression, violence, relationship issues, life threats (suicide and homicide), domestic violence, gambling and financial issues. In the context of this study, dysfunction refers to any state of functioning characterised by psychosocial impairment, where psychopathology is not necessarily present or evident.

AI treatments, according to Fiske *et al.* (2019), may be especially effective at identifying mental health dysfunction at an early stage, interacting with high-risk populations or addressing the concerns of people regarding the social stigma associated with psychotherapy. Patients may occasionally react favourably and productively to the fact that the counterpart is not a real-life therapist.

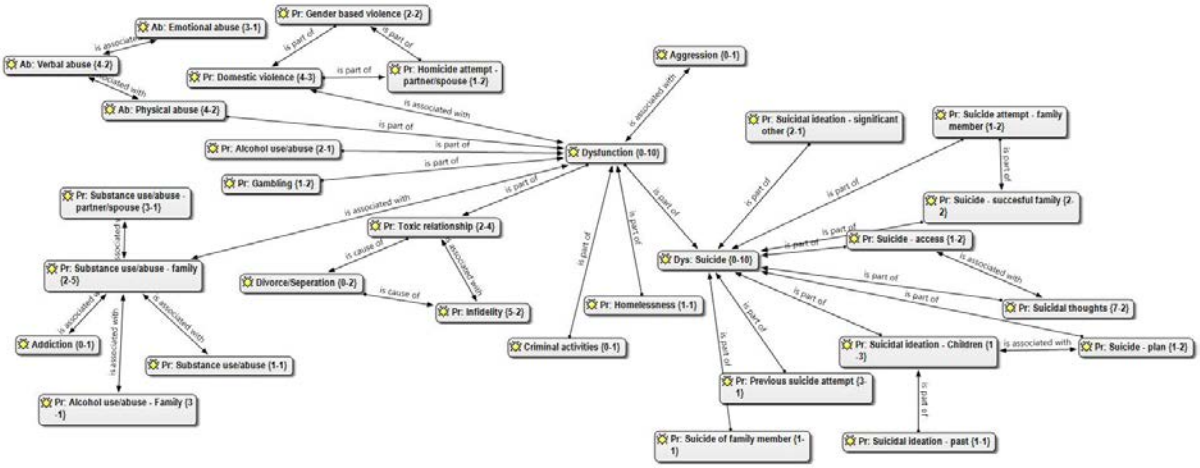


Figure 18. Sub-theme 2.2. Dysfunction network

3.5.2.3. Sub-theme 2.3. Psychological challenges: Uncategorised

In total, 228 codes occurred in Sub-theme 2.3 (Psychological challenges: Uncategorised). Codes that co-occurred include Attention deficit disorder {0-1}; Body image {0-3}; Existential anxiety/crisis {0-1}; Identity {0-1}; Imposter syndrome {0-1}; Memory loss {0-1}; "Mental breakdown" {2-1}; Abandonment issues {1-4}; Adjustment to life stages {1-1}; Anger {8-2}; Attachment style {1-2}; Burden of support {3-2}; Burnout {1-1}; Low mood {1-3}; Low on energy {1-7}; Self-blame {3-0}; Self-doubt {2-2}; Self-harming thoughts {1-2}; Self-resentment {4-4}; Stress {19-8}; Suicidal ideation - past {1-5}; Suicidal thoughts {7-7}; Suicide - access {1-4}; Suicide - plan {1-4}; Anxiety {27-4}; Bipolar mood disorder {1-1}; Complicated grief {1-1}; Depression {13-3}; Learning disorder - dependent {1-1}; Panic attack {3-3}; Phobia {1-1}; Postnatal depression {1-1}; Psych challenges - significant other {1-1}; Substance induced {1-1}; Psychological challenges: Uncategorised {0-30}; Self-esteem {0-3}; Sleep deprivation {0-1}; Social anxiety {0-3}, and Social isolation {0-1}. Excerpt 3 illustrates instances of interactions relating to Sub-theme 2.3 (Psychological challenges: Uncategorised)**.**

2(445) – User: I am also struggling to concentrate and get any work done. I can't focus long enough to get anything meaningful done

2(504) – User: so that is what it has come to. life is so screwed. i am so screwed that we must accept things as they are.

2(76) – User: Because I am hurting alone. You guys dont ever think about me. The last time you guys think about me, is when my last message is sent. Not a minute after that. I hurt while the people I hurt for never even think about me

2(87) – User: I am also feeling like a failure regarding my ability to procreate, given that I am still young and that this was a scenario I only thought I would deal with after the age of 35.

2(197) – User: I was. I explained that when something happens to me emotionally, I unintentionally close up. I'm not a very verbal person about my emotions. some days I look happy but on the inside I'm dying of sad or angry…

Excerpt 3. Psychological challenges: Uncategorised

Sub-theme 2.3 encompasses an array of suspected-, possible- or emerging psychological challenges that cannot be clearly defined or categorised as psychological disorders but are characterised by notable symptom severity and impairment (Michel *et al.*, 2016). The analysed live chat transcript illustrates these as displayed in Figure 19. The live chat transcripts illustrate the complex and multifaceted nature of psychosocial challenges that users sought assistance for in their engagements. The depth of emotional distress and need for understanding, empathy and therapeutic intervention was evident in the analysed data. A study undertaken by Fulmer *et al.* (2018) concludes that AI has the potential to be a practical and affordable therapeutic tool, as results have yielded a significant difference between the control group and the experimental group. In their study, the experimental group indicates a reduction in depressive symptoms (P=0.3) as measured by the PHQ-9.

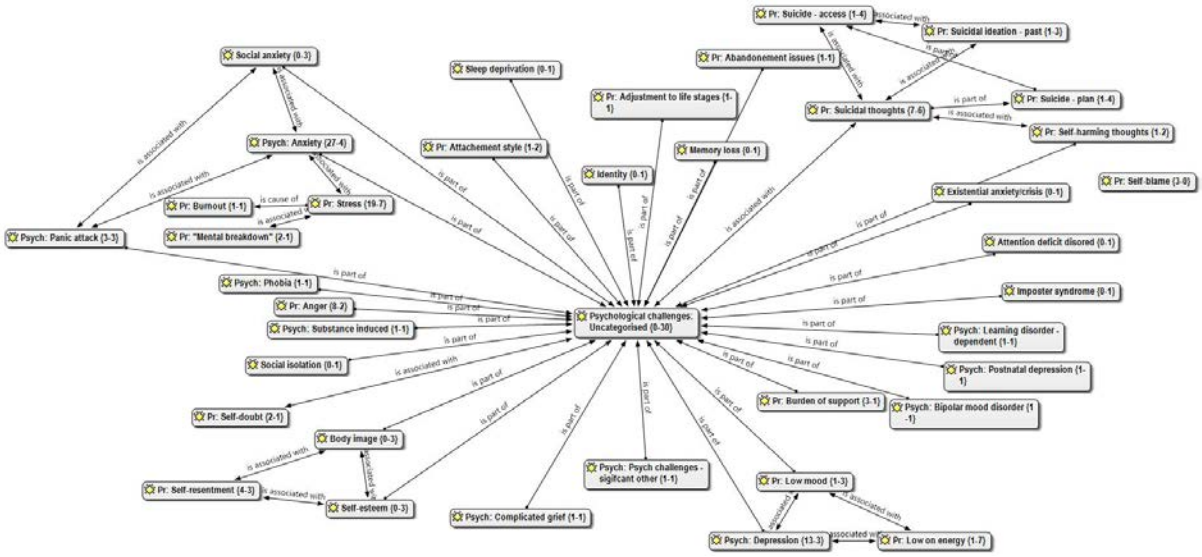


Figure 19. Sub-theme 2.3. Psychological challenges: Uncategorised network

Topoco *et al.* (2019) and Marcelle *et al.* (2019) have investigated treatment consisting of Web-based CBT self-help material and weekly therapist chat sessions, blended with treatment for depression. The intervention attracted persons experiencing depression who needed mental healthcare. The study yielded positive rates in treatment completion rates and substantial improvement in depression symptoms. The outcomes demonstrate that a text-based blended approach can administer CBT in accordance with the pressing need for easily accessible behavioural interventions. In addition, Hull *et al.* (2020) states that improvement rates are in line

with face-to-face therapy, pointing to the viability of two-way messaging as a delivery method for psychotherapy.

3.5.2.4. Sub-theme 2.4. Emotional wellbeing

In total, 196 codes occurred in Sub-theme 2.4 (Emotional wellbeing). Codes that co-occurred include Emotional abuse {3-4}; Verbal abuse {4-3}; Chronic pain {0-1}; Emotional Wellbeing {0-16}; Burden of support {3-2}; Delayed grief {1-1}; Despondency {1-3}; Domestic violence {4-5}; Emotional detachment {1-3}; Emotional disconnection - partner/spouse {1-4}; Emotional regulation {5-3}; Feeling "down" {4-3}; Feeling "numb" {4-2}; Feeling like a burden {1-3}; Feeling overwhelmed {11-1}; Feeling pathetic {1-3}; Feeling worthless {1-3}; Feelings of guilt {5-3}; Isolation {1-2}; Loss and bereavement {15-4}; Negative thoughts {1-3}; Ruminating {1-1}; Self-doubt {2-2}; Self-esteem {2-3}; Self-resentment {4-4}; Spousal manipulation {5-3}; Stress {19-8}; Struggling to adapt after loss {1-1}; Traumatic loss {1-1}; Unhappy {1-2}, and Work-life balance {2-1}. Excerpt 4 illustrates instances of interactions relating to Sub-theme 2.4 (Emotional wellbeing).

2(5) – User: I am one who's has all these compressed emotions I have been suppressed from a young age and it followed me into my marriage

2(29) – User: My heart is feeling heavy. Like I need to cry. I am feeling a bit sad that I am not living my life to it's fullest.

2(60) – User: Hi. I need help. Feeling overwhelmed and it is quite debilitating. My mind does not switch off and I am struggling with insomnia also.

2(82) – User: it's thrown me off everything completely, it's all kind of overwhelming. It's been about a month since I got the court order and the children are rumored to be leaving this month. Little things bring me to my knees at times and I've been taking those moments as they come. But it hasn' t gotten any easier I don' t know if I'm managing.

2(100) – User: I have been taking care of my parents and siblings for a few years with covid 19 I can' t take care of everyone anymore and I have told my parents about it. but now it seems they don' t want to understand that and they have told me to look for my own place and my siblings as well since we have staying at a family house. I'm feeing a lot of emotions because there' s no peace in the my mother is constantly shouting and saying a lot of heartbreaking things

Excerpt 4. Emotional wellbeing

There are three aspects of mental health listed by Galderisi (2015). These are emotional wellbeing, psychological wellbeing, and social wellbeing. Happiness, interest in life, and satisfaction are examples of emotional wellbeing. Psychological wellbeing comprises loving most of one's personality, being adept at handling everyday obligations, having positive relationships with other people, and being content with one's own life. In addition, having something to offer

society, feeling like a part of a community, thinking that society is improving for everyone, and feeling that how society operates makes sense to them are all aspects of social wellbeing, which is defined as positive functioning. The focus of emotional wellbeing is on the users' emotional states and the difficulties they express in maintaining emotional equilibrium and wellbeing. The live chat transcripts indicate numerous instances of people discussing their emotional issues and asking for assistance in dealing with them, as indicated in Figure 20.

Koutsouleris *et al.* (2022) and Tutun *et al.* (2023) believe that advances in AI promise to usher in a new era in which mental health challenges can be effectively detected. Additionally, individuals can receive immediate, precise, and contextualised care; AI advances also empower mental healthcare professionals to make better judgements as digital technology enhances mental healthcare.

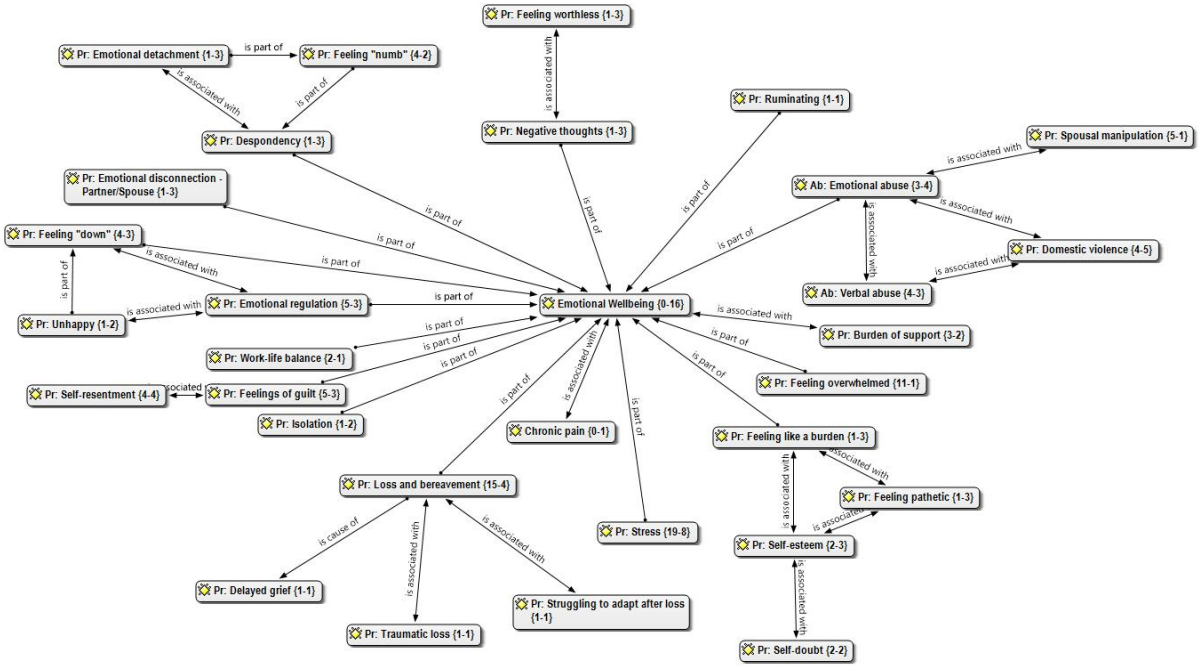


Figure 20. Sub-theme 2.4. Emotional wellbeing network

3.5.2.5. Sub-theme 2.5. Career and work-related: General

In total, 174 codes occurred in Sub-theme 2.5 (Career and work-related: General). Co-occurring codes include Career advancement/Prospects {0-4}; Career aspirations {0-2}; Career planning {0-4}; Career uncertainty {0-2}; Fatigue {0-1}; Overworked {0-2}; Professional development {0-5}; Salary {0-1}; Time management {0-2}; Work ethics {0-1}; Work/life balance {0-1}; Workload management {0-2}; Workplace alcohol use {0-2}; Workplace: Legal {0-2}; Workplace communication {0-1}; Workplace pressure {0-1}; Abuse at work {3-2}; Adjustment - undefined {1-1}; Anxiety at work {1-1}; Application for alternative position {1-1}; Burnout {2-1}; Career path {1-4}; Conflict with manager {2-1}; Contemplating resignation {4-1}; Cultural fit {1-1};

Disability/Incapacity {1-1}; Failing in role {1-2}; Favouritism {1-1}; Feeling redundant {1-2}; Frustration {1-1}; Job security {1-1}; Lack of stimulation {1-1}; Management style {1-1}; Missing deadlines {1-2}; Mistakes at work {1-2}; Need for upskilling {1-0}; New role {2-2}; Not taking leave {1-1}; Overwhelmed at work {1-1}; Performance appraisals {3-1}; Personal stressors influence occupational performance {1-1}; Procrastination {1-1}; Racial profiling {1-2}; Remain focused despite personal issues {1-1}; Restructuring {1-2}; Stagnation in job {2-1}; Transfer {1-1}; Working in a man's world {1-2}; Workplace bullying {1-2}; Workplace dissatisfaction {6-0}; Workplace stress {3-3}, and Written warning {2-1}. Excerpt 5 illustrates interactions relating to Sub-theme 2.5 (Career and work-related: General).

2(60) – User: I am good at managing people and product delivery. The whole landscape has changed though ... And that is very challenging. Not sufficient resources, support functions not optimal, diluted focus ... No consistency with roles and responsibilities across the bank. Still a work in pogress

2(95) – User: since I started working I have also felt like I'm constantly competing. I know I have to prove myself but to some extent it feels its too much

2(171) – User: I am expected to just help everyone around here , which I don' t mind about who helps me when I'm not around ? I have raised this before

2(222) – User: my job is not fulfilling me I have been on the same role for over four years now it's not stimulating it's boring and my nett is decreasing instead of increasing because of medical aid costs and compulsory deductions I feel stagnant my current role is not serving me I am just working because working but with the cost of living going up I don't enjoy my salary I feel like a volunteer I'm overqualified for the role I am and thought I put it ally effort and the company is aware of this and I'm furtheringy studies there has been a resistance toy growth so many

2(421) – User: Currently I am dealing with 2 grievances involving her. The one employees in my team lodged against her and the one she lodged as her performance rating dropped. Yesterday I was frantically called about another incident again involving her with another senior manager. It never stops

Excerpt 5. Career and work-related: General

Sub-theme 2.5 (Figure 21) refers to any direct or indirect challenges that the users experienced within their place of employment. These issues range from work discontent to the influence of workplace frustrations on the user's personal life. The contemporary workplace is turbulent and stressful. Career actors are increasingly expected to be responsible for their professional growth as traditional linear organisational careers continue to give way to self-directed, values-driven, and boundaryless modern careers (Lyons *et al.*, 2015). According to Argyropoulou and Kaliris (2018), the current economic crisis has significantly reduced the margin for error while dramatically increasing uncertainty to a very considerable extent and making career planning decisions difficult. These assertions are supported by the reasons for which individuals connected to live chat to converse about professional and work-related difficulties. Magogodi's (2017) study

reveals how certain aspects of text-based counselling act as facilitators, enabling interventions to be effectively integrated and resulting in a high success rate for resolving workplace issues.

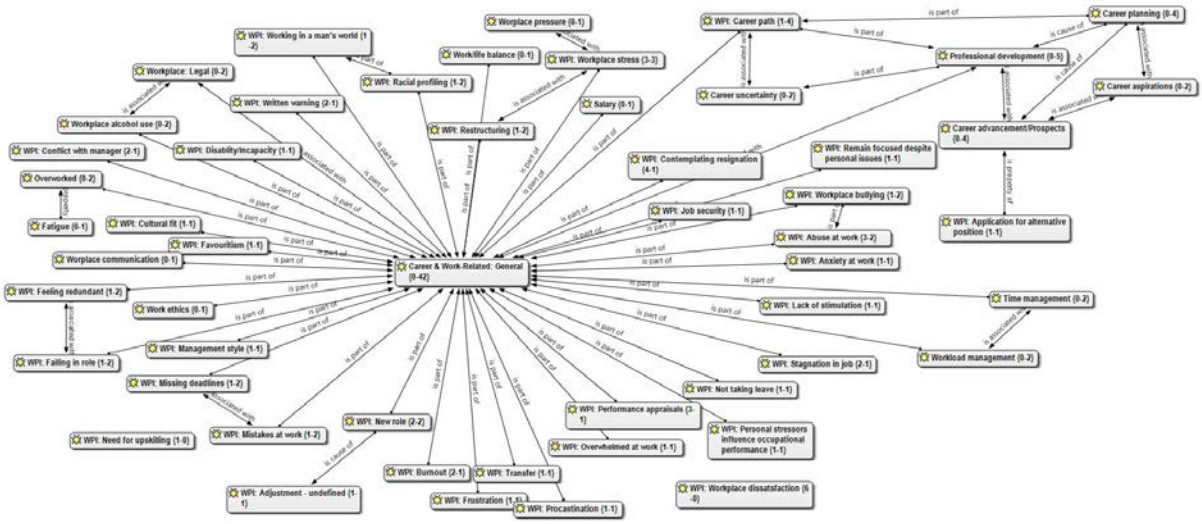


Figure 21. Sub-theme 2.5. Career and work-related: General network

3.5.2.6. Sub-theme 2.6. Interpersonal relationships

In total, 96 codes occurred in Sub-theme 2.6 (Interpersonal relationships). Co-occurring codes include Changes in roles {0-1}; Intimacy {0-1}; Marriage scepticism {0-1}; Non-conventional relationships {0-1}; Office romance {0-1}; Open relationship {0-1}; Anger - partner/spouse {2-4}; Anger towards partner/spouse {1-6}; Children from previous relationship/marriage {1-2}; Contemplating leaving relationship/marriage {1-4}; Emotional disconnection - partner/spouse {1-4}; Facebook - Partner {1-3}; Failed relationships {2-2}; Fear of separation {2-3}; Fear of upsetting partner/spouse {1-1}; Frustrated in relationship {1-1}; Infidelity {5-6}; Infidelity - user {3-1}; Interpersonal conflict - partner/spouse {11-5}; Interpersonal relationship challenges - Developing relationship {1-2}; Interpersonal relationship challenges - partner/spouse {19-1}; Lying - self {1-2}; New family composition {1-2}; Not ready for marriage {1-1}; Partner/Spouse abuse {1-1}; Partner/Spouse going through a phone {1-1}; Possible relationship {1-3}; Pressure from parents to divorce {1-1}; Rejection - romantic {1-2}; Relationship - trust issues {4-2}; Relationship/marriage undefined {5-36}; Separation/Divorce {17-3}; Sexual dissatisfaction in relationship {3-1}; Spousal manipulation {5-3}; Substance use/abuse - partner/spouse {3-2}; Toxic relationship {2-7}; Untruthful - partner/spouse {1-1}; Relationship boundaries {0-1}; Relationship building {0-1}; Relationship preferences {0-1}; Romantic interest {0-2}; Separation recovery {0-2}, and Unrequited love {0-2}. Excerpt 6 illustrates interactions relating to Sub-theme 2.6 (Interpersonal relationships).

2(44) – User: my husband and I are going through a rough patch he was planning on telling me he was leaving but I had a breakdown before he did and he has stayed

2(122) – User: I'm so stressed. my partner has changed into someone I don't know. his behaviour has been causing fights between us and we are just drifting apart. my biggest problem is he is neglecting his responsibility to take care of the kids financially. He sent his family (mother, father and brother) money but the kids don't have clothes, we don't have food, we haven't paid rent nor the accounts. next week they are supposed to go to creche, he's not giving me the money for that. I really don't know what I'm going to do, I can't even sleep.

2(152) – User: I would like to be sure if the marriage is still viable and if we can both acknowledge our faults and the role we play in what causes the problems in the marriage. and even if the marriage cannot survive then for us to become better people for the sake of the kids. I would also like for the kids to get counseling as this trauma and uncertainty has also affected them.

2(225) – User: my mother has never liked my husband and she confused my 10 year old daughter and turned her against my husband

2(294) - User it's clear to me that there seems to be a problem within the marriage but he only creates issues when he drinks and he drinks all the time. when in a normal frame of mind. he hardly speaks to me. we have no intimate relationship for over a year. I don't think he has someone else because he is always at hme

Excerpt 6. Interpersonal relationships

The investigation of interpersonal relationship dynamics constitutes a dedicated and independent field of study. However, within the context of this study, the scope does not encompass an in-depth analysis of relationship challenges and the corresponding strategies for resolution. Throughout this study, individualised counselling services were provided to the users to address their unique concerns. It is important to emphasise that the approach adopted in this research did not involve the provision of group therapeutic interventions to the participants. The focus remained on understanding the dynamics of live chat interactions and the efficacy of personalised counselling in addressing the complexities of interpersonal relationships, as displayed in Figure 22.

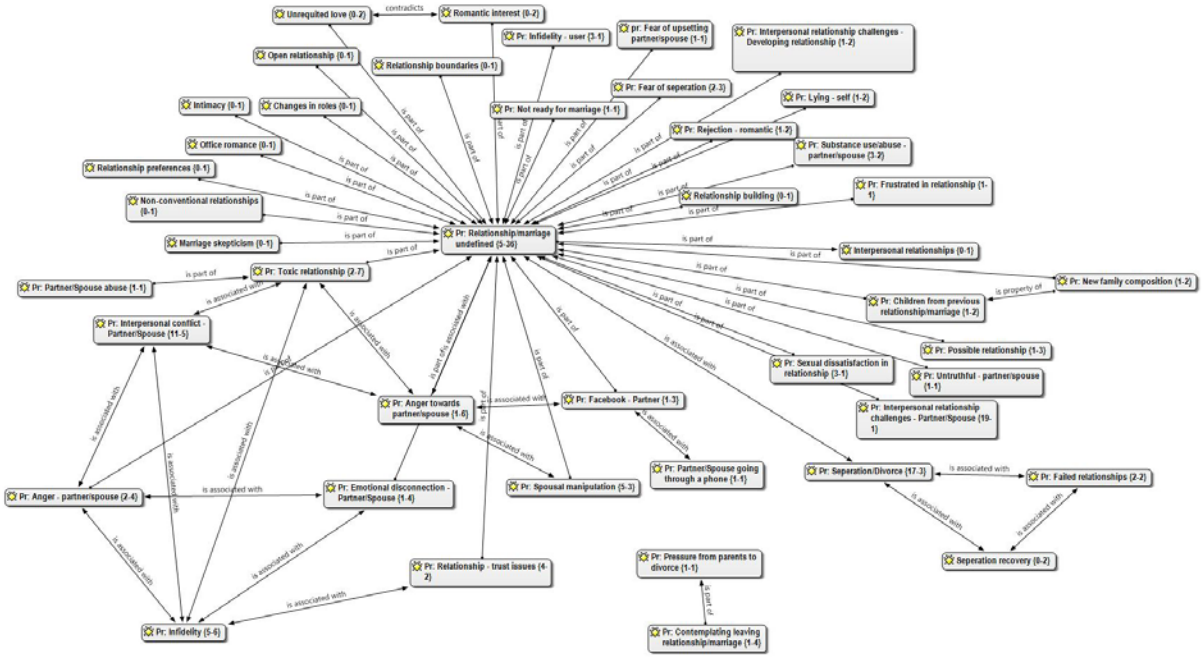


Figure 22. Sub-theme 2.6. Interpersonal relationships network

3.5.2.7. Sub-theme 2.7. Family dynamics

In total, 101 codes occurred in Sub-theme 2.7 (Family dynamics). Co-occurring codes include Blended families {0-1}; Co-parenting {0-2}; Family enmeshment {0-1}; Family finances {0-3}; Family planning {0-1}; Family secrets {0-1}; Fatherhood {0-3}; Generational conflict {0-1}; Motherhood {0-2}; Parental relationship {0-1}; Child development {1-2}; Abandonment issues {1-4}; Absent father {4-3}; Anger towards family {1-1}; Childcare - aggression {1-1}; Childcare - anger {2-3}; Childcare - discipline {3-3}; Childcare - general negative behaviour {1-6}; Childcare - respect {1-1}; Childcare - School issues {1-2}; Childcare - suspension {1-4}; Childcare - tantrums {1-1}; Childcare: Corporal punishment {1-2}; Childhood abandonment by father {1-5}; Divorced parents {1-1}; Family dynamics {2-32}; Finances - retrenchment - breadwinner {1-3}; Financial constraint - support extended family {1-3}; Interpersonal conflict - family {6-2}; Primary place of residence {1-1}; Single parent {4-3}; Suicide - successful family {2-4}; Suicide attempt - family member {1-4}; Suicide of family member {1-3}, and Traditional gender roles {0-1}. Excerpt 7 illustrates interactions relating to Sub-theme 2.7 (Family dynamics) (Figure 23).

> *2(5) – User: My parents are having a hard time my dad is self employed and his health has recently taken a bad turn but is now recovering and my mum is also medically boarded my parents have made bad decisions for themselves which is no affecting my marriage*
>
> *2(110) – User: I would like to make contact with my father whom I've been searching for all my life. I have got the information from my mother and I know where he is currently*

2(135) – Use: My mother is taking out that she knows of my father' s affair on the side. I spoke to him that I do not like being spoken to like that. He responded that I made such a small thing a big thing. I do not use foul language

2(142) – User: Having a child has proven to one of the hardest things, especially without support from family and the father not being present all together Agent Tell me more about what you have experienced being a parent.

2(243) – User: now that we are no longer in speaking terms there tension in our family. We no longer spend a lot of time as a family cause when she's drunk she always will pass rude remarks and hurt people.

Excerpt 7. Family dynamics

Augar and Zeleznikow (2014) believe that participation in online text-based support and counselling in enhanced family dispute resolution through the use of technology is often essential and overlooked. Users may benefit from the inclusion of online counselling to support networks by being enabled to receive personalised input in a more private online environment.

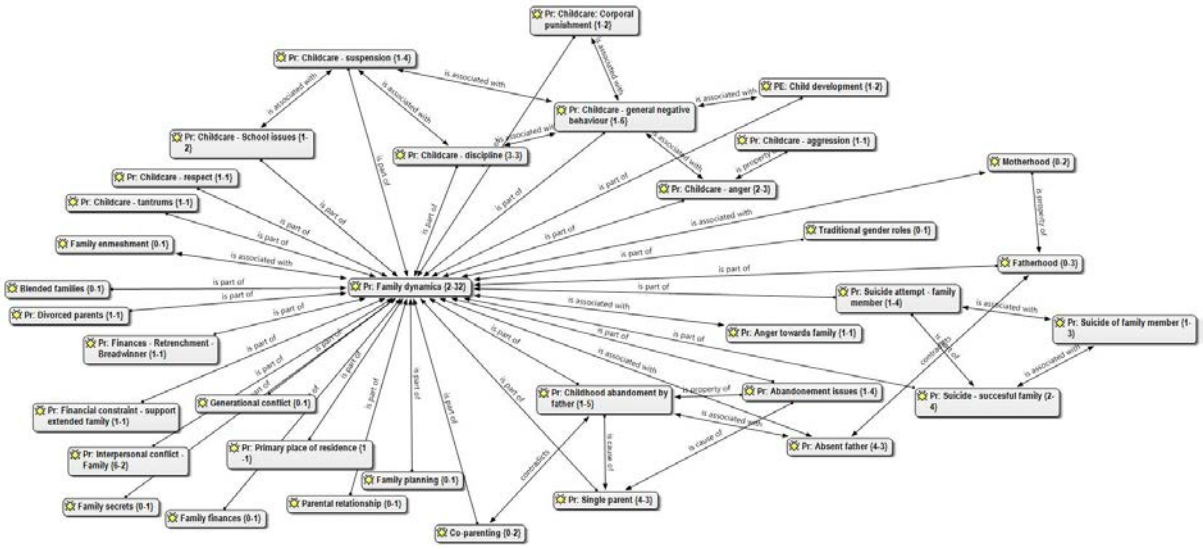


Figure 23. Sub-theme 2.7. Family dynamics network

3.5.2.8. Sub-theme 2.8. Request for help/support

Users' requests for help/support co-occurred 193 times, with the most prominent requests relating to Advice seeking {14-3}; Information seeking {0-1}; Need for clarification {0-0}; Problem acknowledgement {0-2}; Request feedback {0-0}; Seeking external validation {0-0}; Alternative coping strategies {1-0}; Child counselling {3-1}; Clarity {2-1}; Diverse {2-0}; Need for ongoing support {1-1}; Need for resources {1-0}; Parental guidance/skills {3-1}; Prepare for divorce {1-0}; Psychiatric assistance {1-0}; Seek guidance {2-3}; Support {1-1}; Tools to manage anxiety/stress {1-1}, and Understand anxiety {1-1}. Excerpt 8 illustrates interactions whereby users requested help/support. These relationships are displayed in Figure 24.

2(19) – User: I am struggling with marital issues, and its definitely affecting my mental health. We have a certain issue that we've been having and I felt the need to speak to a counselor and he should as well in order for us to know the best way to move forward […] I personally am more comfortable with face to face or telephonic because I'm very much a talker

2(40) – User: I've been reflecting the past 2 weeks and I decided to at least speak with someone and make use of ICAS

2(47) – User: do u think that there is someone that u can recommend that he chats to through this platform? or is it done via whatsapp like this? not sure how it works

2(541) – User: hence I said, it's like I am fighting a devil that I don't know […] please advise of someone in your profession

2(415) - User What is the process to follow for thhe intake/assessment sessions and then the face to face counselling.

Excerpt 8. Request for help/support

The method through which users sought help differed from person to person. While some users promptly explained the situation they needed assistance with, others preferred to request help by letting the agent inquire further to collaboratively construct a detailed problem description. Gibson and Cartwright's (2014) study proposes that text messages on mobile phones could align effectively with the preferences and interests of (especially young) individuals. The elements that participants considered valuable in text-based counselling resemble those found in previous research (Binder *et al.*, 2011; Gibson & Cartwright 2013) on people's counselling experiences. This resemblance is evident in the sense that such aspects seemed to address their desire for control and independence.

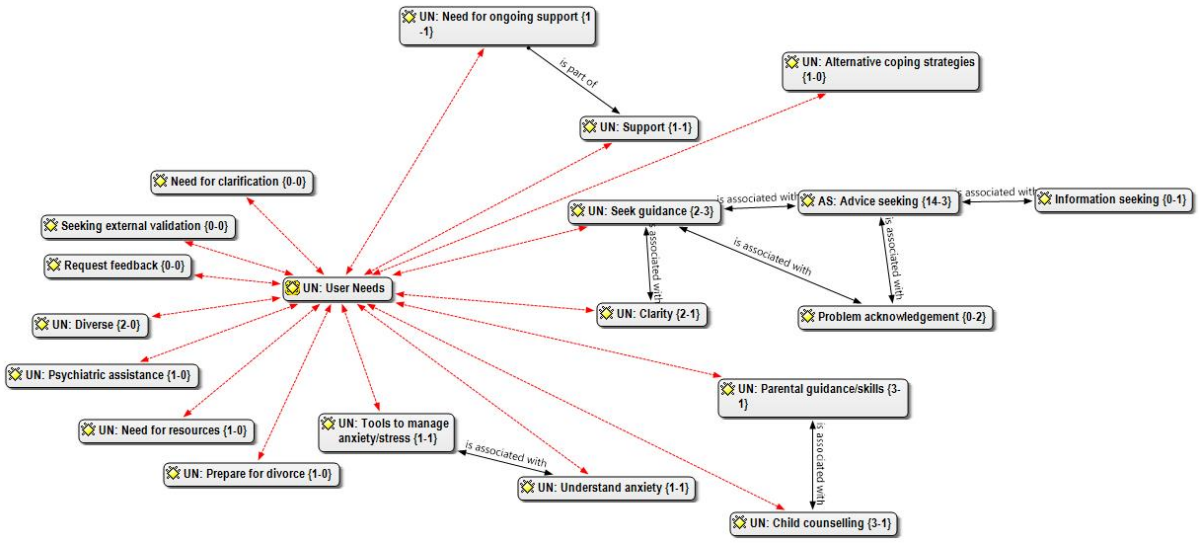


Figure 24. Sub-theme 2.8. Request for help/support network

3.5.2.9. Sub-theme 2.9. Suicidal ideation

Suicide or suicidal ideation, which is viewed as a risk code, occurred seven times in the analysed transcripts. Co-occurring codes include Suicide {0-12}; Previous suicide attempt {3-2}; Suicidal ideation - past {1-5}; Suicidal thoughts {7-7}; Suicide - access {1-4}; Suicide - plan {1-4}, and Suicidal behaviour {0-2}. Excerpt 9 illustrates interactions relating to Sub-theme 2.9 (Suicidal ideation). The relationships are displayed in Figure 25.

2(171) – User: I know I wouldn' t attempt to commit suicide this time , but it' s a feeling that I feel , like I just ask myself if it would be better if weren' t here
2(265) – User: i'm also worried about overdosing and committing suicide because i've been thinking about it for a while not
2(504) – User: i think i might be suicidal

Excerpt 9. Suicidal ideation

Mathieu *et al*.'s (2021) study has found that developments in the mode of communication, specifically online communication and counselling, have become an emerging trend. This online communication relates to emotional support and suicide prevention for (particularly) youth, which has proven to be effective in increasing users' resilience and enabling timely identification, intervention and management of psychological issues and ultimately reducing suicidal ideation (Slomski, 2015).

At an individual level, predictive analytics play a crucial role in recognising individuals facing crises. They enable timely interventions through emotional support, access to crisis and psychoeducational resources, as well as prompt alerts for emergency assistance. On a broader scale, these algorithms can identify vulnerable groups and areas with high suicide rates, thereby aiding in the allocation of resources, policy improvements, and advocacy endeavours. AI has also found utility in supporting the clinical management of suicide, encompassing diagnostics, evaluation, medication administration, and behavioural therapy. The integration of AI into suicide care offers numerous advantages, such as providing a time- and resource-efficient alternative to clinician-based methods, adaptability to diverse environments and demographics, and its applicability in remote locations with limited access to mental healthcare services (Fonseka *et al.*, 2019).

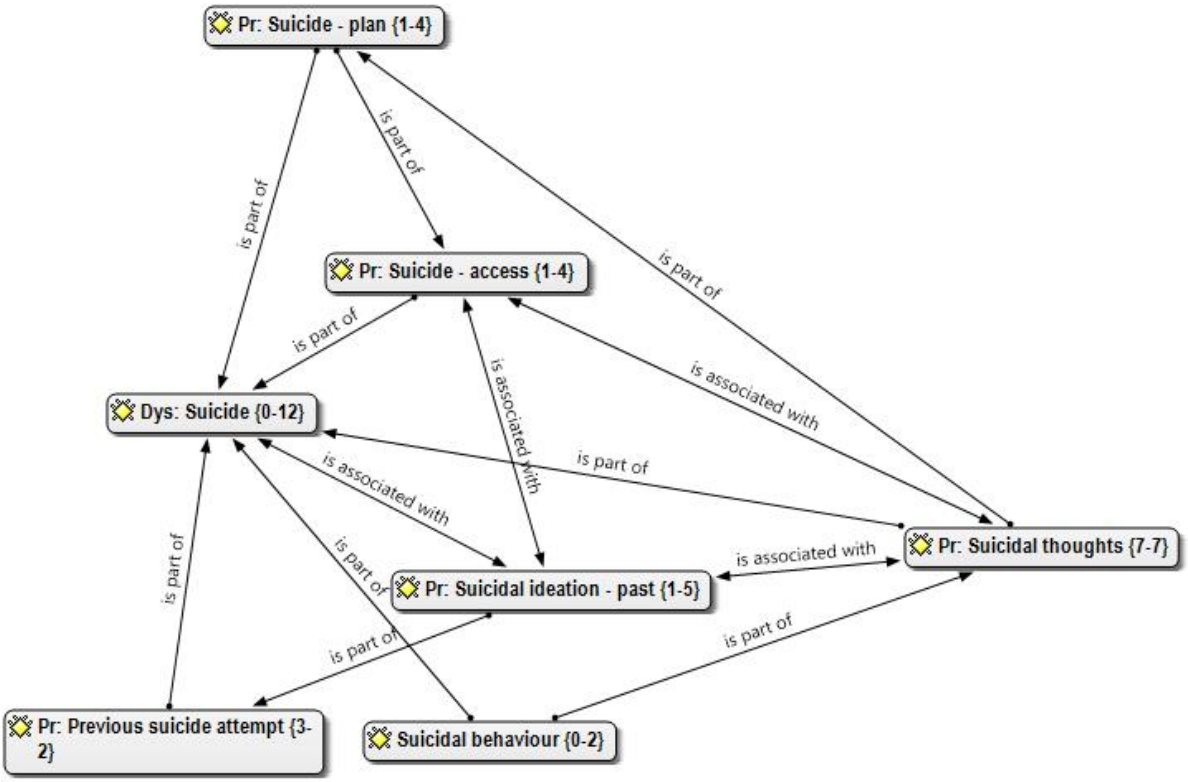


Figure 25. Sub-theme 2.9. Suicidal ideation network

3.5.2.10. Sub-theme 2.10. Diverse challenges/issues/problems

Users connected to live chat with a variety of issues and challenges which may not always fit into a specific category. Thus, Sub-theme 2.10 was not quantified. These diverse problems often fell within minority categories and could not be grouped into a predefined code group or code family. The significance of this theme is that one can never predict or be fully prepared for the spectrum of issues that a live chat user might connect for assistance. Excerpt 10 illustrates interactions relating to Sub-theme 2.10 (Diverse challenges/issues/problems).

2(359) – User: I am in a relationship (love related relationship) with a guy and it's been 1 year and 6 months .... Things were great when we started but now I feel like I'm in an emotionally abusive relationship that I can't leave because when I leave he comes and begs me back and I fall for it ..... Now I can't even express my feelings or emotions in the relationship because he is always going through a lot so I have to put my feelings aside and look out for his .... It's not just him ... Family

2(351) – User: I want us to have a second child now. I recently found out I have some issues with my uterus so I'm running out of time

2(67) – User: I'm even scared to go test cause I don't think there is someone that can use protection with someone that, his been sleeping with for four years.

2(484) – User: I was diagnosed with cancer in March 2021. I'm in remission. lost my mom to ovarian cancer Feb 2022. I was on disability leave for a year

2(40) – User: and again august my close cousin committed suicide ?? just after my birthday still i went on with my life

Excerpt 10. Diverse challenges/issues/problems

3.5.3. Theme 3: Counsellor approaches

Counsellor approaches to live chat sessions were analysed. The analysis focused on how counsellors approach the specific needs as described by the service users (Figure 26). There were three overarching themes, namely advice and support, psychological approach and therapeutic approach, which were identified (as displayed in Table 7).

Table 7. Counsellor approaches

Code group	Count of code group
Advice & Support	63
Psychological approach	39
Therapeutic approach	156
Grand Total	**258**

Advice and support in the context of this study refer to actionable advice which is given to the user by the agent. For example, *[2(363) - Agent: It is a difficult situation to be in and these are big decisions to make about your life, maybe it could help to write down your thoughts to sort them out and process them].*

Psychological approach within the context of this study refers to the psychological paradigm from which the agent conducts their intervention. *For example, [2(376) – Agent: Tell more what would need to change for you to have better experience at work? what would need to be different?].* This is an example of CBT.

Therapeutic approach in the context of this study refers to the therapeutic skills that an agent would draw upon to supplement the therapeutic paradigm from which they work. For example, *[2(29) – Agent: We all do this, but it is something that we need to be aware of and change.]* This is an example of normalising.

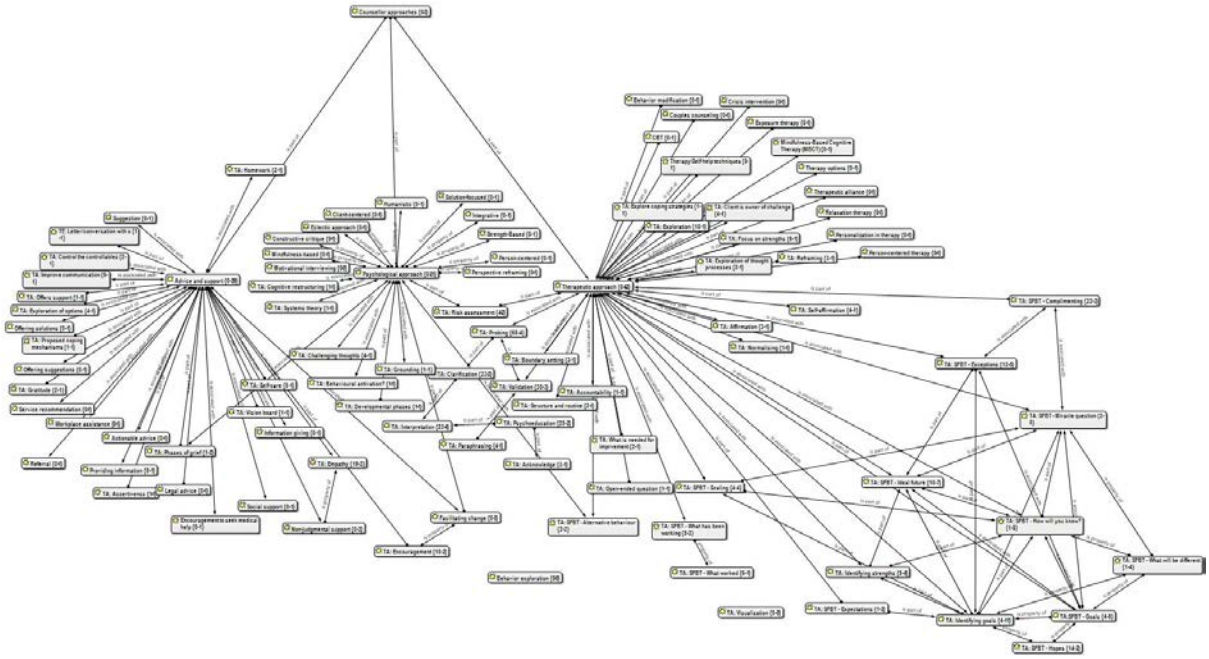


Figure 26. Counsellor approaches

3.5.3.1. Sub-theme 3.1. Advice and support

In total, 63 themes relating to Sub-theme 3.1 (Advice and support) were identified. Co-occurring themes include Actionable advice {0-4}; Advice and support {0-38}; Couple's counselling explored {2-3}; Couple's counselling recommended {6-6}; Exploring couple's counselling {1-2}; Creativity {2-4}; Journaling {8-3}; Letter writing {1-4}; Encouragement to seek medical help {0-1}; Facilitating change {0-4}; Information giving {0-4}; Legal advice {0-1}; Offering solutions {0-1}; Offering suggestions {0-5}; Providing information {0-5}; Referral {0-3}; Service recommendation {0-3}; Social support {0-1}; Suggestion {0-2}; Acknowledge {3-2}; Affirmation {3-2}; Assertiveness {1-2}; Control the controllables {2-2}; Encouragement {10-3}; Exploration of options {4-2}; Gratitude {2-2}; Homework {2-2}; Improve communication {6-2}; Offers support {1-2}; Proposed coping mechanisms {1-2}; Self-care {5-2}; Vision board {1-3}; Letter/conversation with x {1-2}, and Workplace assistance {0-1}. Excerpt 11 illustrates instances where Sub-theme 3.1 was evident.

2(171) – Agent: That suggests some creativity on your part .....which you won't really find in the accounting world.

2(222) – Agent: Try and see if you can also express your feelings in a tangible or creative way - Write about your loss in a journal. If you' ve lost a loved one, write a letter saying the things you never got to say; make a scrapbook or photo album celebrating the person' s life; or get involved

2(421) – Agent: I see. Boundaries are crucial in the workplace. And I understand that it might feel easier said than done at this point in time . . But [name], I want to encourage you to be assertive in establishing clear boundaries. Ensure that the conversations and interactions between the two of you are strictly work related. Perhaps it might be a good idea to also keep record of every interaction.

2(185) – Agent: Look yourself in the mirror and repeat words of affirmation to yourself . . continuously tell yourself that you are enough .. You are strong and you deserve someone who will love you without uncertainty . Remind yourself of how much you love you.

2(440) – Agent: For starters you need to give up berating yourself for everything you aren' t. - Being kind to yourself in thoughts, words and actions is as important as being kind to others. Extend yourself this courtesy. Love yourself - your real self. Work through your fears (dive deep), your insecurities (speak honestly and loudly), and your anger (scream into the pillow - not into the mirror, nor the people you care about; they don' t deserve it.) Instead of hurting yourself by hiding from your problems, help yourself grow beyond them. That' s what self-care is all about.

2(49) – Agent: at this point we need start looking at what you have in control.

2(14) – Agent: It's amazing that you are able to still take care of your personal life during this time, it takes a lot of strength and courage, you can be

proud of yourself. Has there been a time at work when things were a little better, even just for a brief moment?.

2(171) - Agent "I have some interesting home work for you. Read up on ""systems theory"". […]Particularly "systems theory and change".

Excerpt 11. Advice and support

Live chat agents adopted a direct approach to offering advice and support, as can be seen in illustrated quotations. Advice and support varied from providing factual verifiable information to professional and/or personal opinions. As displayed in Figure 27, the advice and support followed a common theme, whereby the themes are associated with- and form part of one another.

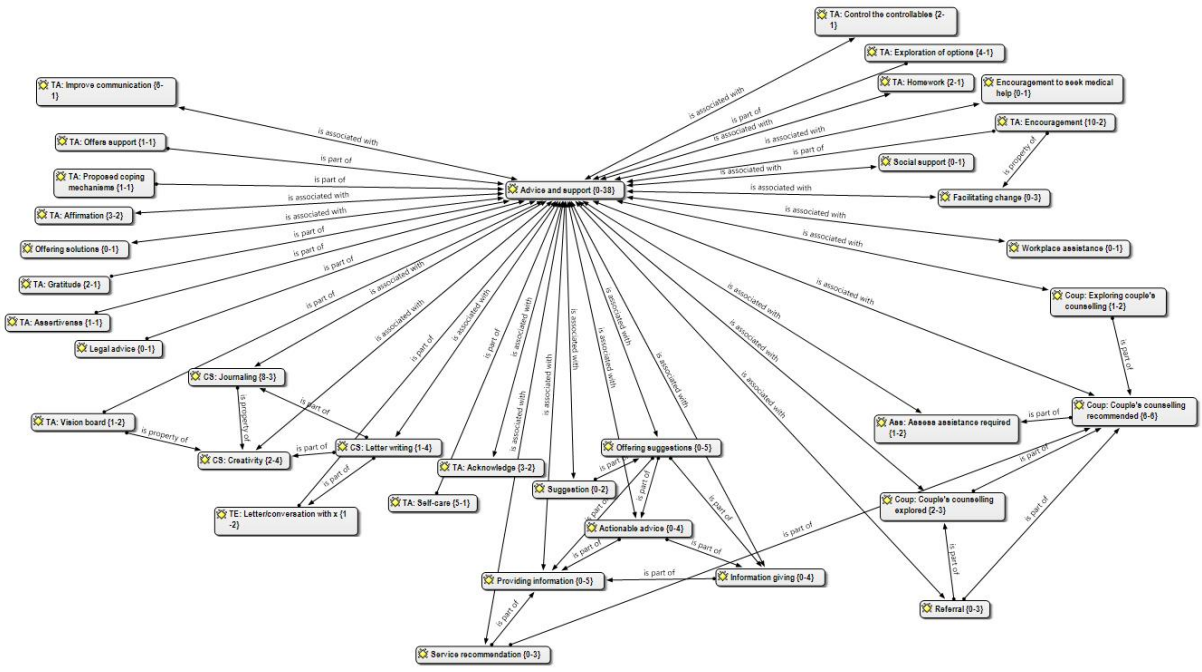


Figure 27. Advice and support

3.5.3.2. Sub-theme 3.2. Psychological approach

In total, 39 themes relating to Sub-theme 3.2 (Psychological approach) were identified. These themes include Client-centred {0-1}; Constructive critique {0-2}; Eclectic approach {0-1}; Humanistic {0-1}; Integrative {0-1}; Mindfulness-based {0-1}; Mindfulness-Based Cognitive Therapy (MBCT) {0-2}; Motivational interviewing {0-1}; Person-{0-1}; Person-centred therapy {0-2}; Personalisation in therapy {0-2}; Perspective reframing {0-1}; Psychological approach {0-25}; Solution-focused {0-7}; Behavioural activation; {1-2}; Challenging thoughts {4-2}; Cognitive restructuring {1-1}; Developmental phases {1-1}; Grounding {1-1}; Risk assessment {4-2}; SFBT - Alternative behaviour {2-2}; SFBT - Complimenting {22-3}; SFBT - Exceptions {12-5}; SFBT - Expectations {1-3}; SFBT - Hopes {14-3}; SFBT - How will you know? {1-9}; SFBT - Ideal future {10-7}; SFBT - Scaling {4-4}; SFBT - What has been working {5-3}; SFBT - What will be different {1-5}; SFBT - What worked {6-2}, and Systems theory {1-1}. Excerpt 12 illustrates the psychological approaches utilised by live chat counsellors.

2(99) – Agent: 1. Become aware of what you are thinking . . 2. Flip it around. Try to see the opposite to that thought. . Let's challenge that thought:

2(72) – Agent: Challenging these negative thoughts is an effective way to reduce the symptoms of social anxiety .. . Step 1: Identify the automatic negative thoughts that underlie your fear of social situations. For example, if you're worried about an upcoming work presentation, the underlying negative thought might be: "I'm going to blow it. Everyone will think I'm completely incompetent." . . Step 2: Analyze and challenge these thoughts. It helps to ask yourself questions about the negative thoughts: "Do I know for sure that I'm going to blow the presentation?" or "Even if I'm nervous, will people necessarily think I'm incompetent?" Through this logical evaluation of your negative thoughts, you can gradually replace them with more realistic and positive ways of looking at social situations that trigger your anxiety.. . It can be incredibly scary to think about why you feel and think the way you do, but understanding the reasons for your anxieties will help lessen their negative impact on your life ..

2 (110) – Agent: Thank you for the information. Moving onto your session, kindly share your best hopes for this session today?.

2(192) – Agent: Grounding Breathing Technique . . I want you to start by closing your eyes and once your eyes are closed I want you to start taking deep breaths in through your nose and then slowly slowly out through your mouth. In through your nose and slowly out through your mouth ... Now while you breath, I want you to start focusing on your other senses and answer the below kinds of questions in your head: . . - What can you feel with your hands? Can you feel the texture of your phone?. - What can you

2(29) – Agent: Good. How will you know when you are happy and okay? What will be different and what will you do differently if you are happy and okay? .

2(40) – Agent: What has been helping you manage all this time? Surely, suppressing your feelings cannot be a healthy way of coping.

2(265) – Agent: On a scale of 0 to 10, if zero was the worst it has been (Your emotional wellbeing), where is it now?.

2(13) – Agent: You are doing well and you should validate your progress thus far , this is not something easy and you are dong well .

2(16) – Agent: Tell me more about your preferred future look like. How would you rather be feeling instead?.

2(197) – Agent: "Suppose that tonight while you were asleep, there was a miracle and this problem that. brought you here today disappeared. The miracle occurs while you are sleeping so you do not immediately know that it has happenedn. "When you wake up, what is the first thing you will notice that will let you know that there. has been a miracle?.

2(25) – Agent: a behavioural activation challenge. how does that work […] which means- doing some activity in your routine is better than not doing anything at all.

Excerpt 12. Psychological approach

SFBT - a paradigm associated with CBT - is Lyra SA's psychological paradigm of choice. The use of CBT and specifically SFBT is evident in the illustrated quotations (Excerpt 12) and the psychological approach network (Figure 28).

Andersson (2016) highlights that CBT has demonstrated efficacy as a therapeutic approach for addressing psychological challenges (especially with regard to depression and anxiety) through online platforms. Moreover, Topooco *et al.* (2018) emphasise that the involvement of and support of counsellors play a pivotal role in ensuring positive outcomes within this therapeutic framework. A meta-analysis by Lim *et al.* (2022) has found that chatbot-delivered psychotherapy significantly improved depressive symptoms by incorporating CBT and PST (problem-solving therapy) into scripts programmed within chatbot software. According to Fernández and Mairal (2017), these scripts contain approaches for cognitive restructuring and behavioural activation that enhance negative cognitions, while methods for problem evaluation and problem orientation lessen perceived problem severity. Within CBT and PST, both the decrease in dysfunctional attitudes and the rise in feelings of control mediated the improvement of depressed symptoms.

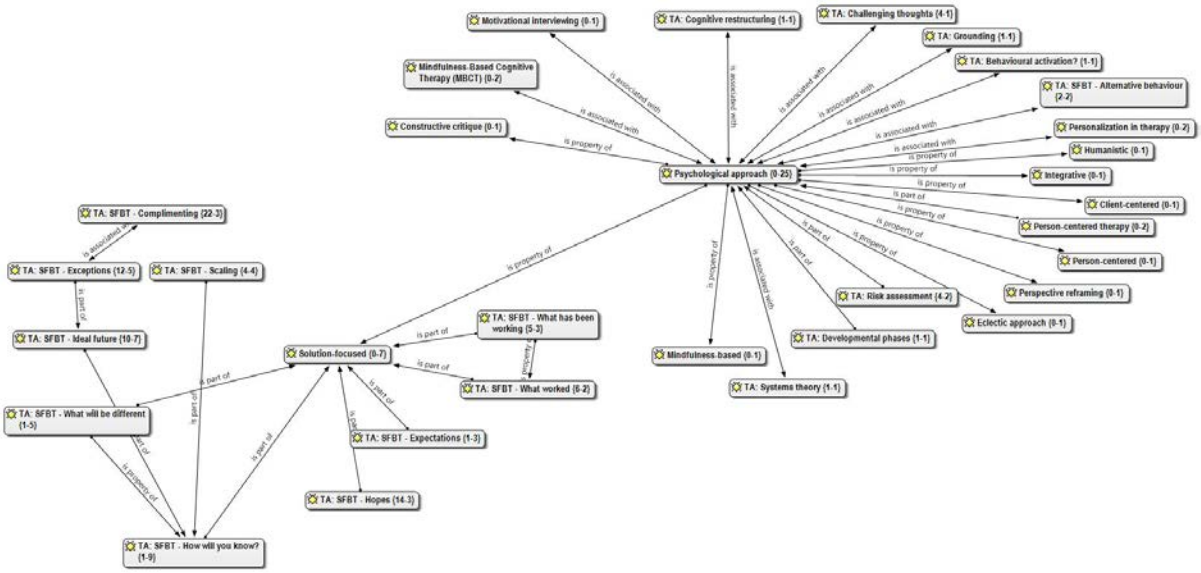


Figure 28. Psychological approach

3.5.3.3. Sub-theme 3.3. Therapeutic approach

The coding of therapeutic approaches serves as a representation of how live chat agents engage in online counselling, encompassing professional counselling techniques utilised by counsellors to provide assistance and support to users. In total, 156 themes relating to Sub-theme 3.3 (Therapeutic approaches) were identified. These themes include Challenging negative thought patterns {1-1}; Constructive critique {0-2}; Crisis intervention {0-1}; No therapeutic approach {1-1}; Facilitating change {0-4}; Affirmation {3-2}; Assertiveness {1-2}; Behavioural activation {1-2}; Boundary setting {3-1}; Challenging thoughts {4-2}; Clarification {23-2}; Client is owner of challenge {4-1}; Control the controllables {2-2}; Distorted thinking patterns {1-1}; Empathy {19-3}; Encouragement {10-3}; Exploration {10-1}; Exploration of options {4-2}; Exploration of thought processes {2-1}; Explore coping strategies {1-1}; Focus on strengths {6-1}; Gratitude {2-2}; Homework {2-2}; Honesty {1-1}; Identifying goals {4-11}; Identifying strengths {5-4}; Improve communication {6-2}; Interpretation {23-3}; Normalising {1-1}; Offers support {1-2}; Open-ended question {1-1}; Paraphrasing {4-1}; Phases of grief {1-3}; Probing {68-4}; Proposed coping mechanisms {1-2}; Psychoeducation {25-2}; Reframing {3-1}; Risk assessment {4-2}; Self-affirmation {4-1}; Self-care {5-2}; Session summary {4-1}; Validation {28-3}, and What is needed for improvement {2-1}. Excerpt 13 and Figure 29 illustrate the use of various therapeutic approaches by live chat counsellors.

2(12) – Agent: You counselling is all about providing you with support, I can hear that you feel very alone at this moment and I understand how harrowing that can be.

2(235) – Agent: When did you start seeing these communication issues? And when you say single mom problems can you elaborate on that further?

2(222) – Agent: With regards to grief i would encourage that you try to face your feelings - You can try to suppress your grief, but you can't avoid it forever. In order to heal, you have to acknowledge the pain. Trying to avoid feelings of sadness and loss only prolongs the grieving process. Unresolved grief can also lead to complications such as depression, anxiety, substance abuse, and health problems.

2(117) – Agent: No I am not at liberty to advise on that. The reason for that only the person that comes for counselling knows their needs and what made them to decide on attending counselling.

2(25) – Agent: from what you are telling me it sounds like you are someone who is able to achieve whatever you put your mind to and i think that goal is very much possible.

2(117) – Agent: 1) Physical Boundaries. 2) Sexual Boundaries. 3) Emotional or

Mental Boundaries. 4) Spiritual or Religious Boundaries. 5) Financial and Material Boundaries. 6) Time Boundaries . .... 7) Non-Negotiable Boundaries .. . Boundaries are what builds or destroys relationships, friendships and families . . It is a way of holding yourself accountable over what you accept and how the next person treats […]For example: saying I do not like how you talk to me , might seem like its a boundary but it is not.

2(152) – Agent: The main sign that a person has bonded with an abuser is that they try to justify or defend the abuse. They may also: . . agree with the abusive person' s reasons for treating them badly. try to cover for the abusive person. argue with or distance themselves from people trying to help, such as friends, family members, or neighbors. become defensive or hostile if someone intervenes and attempts to stop the abuse, such as a bystander or police officer . . Based on the above, I do not think that it is trauma bonding. You do not accept his behavior as a definition of you and your worth. It is important to realize where you put the blame.

Excerpt 13. Therapeutic approach

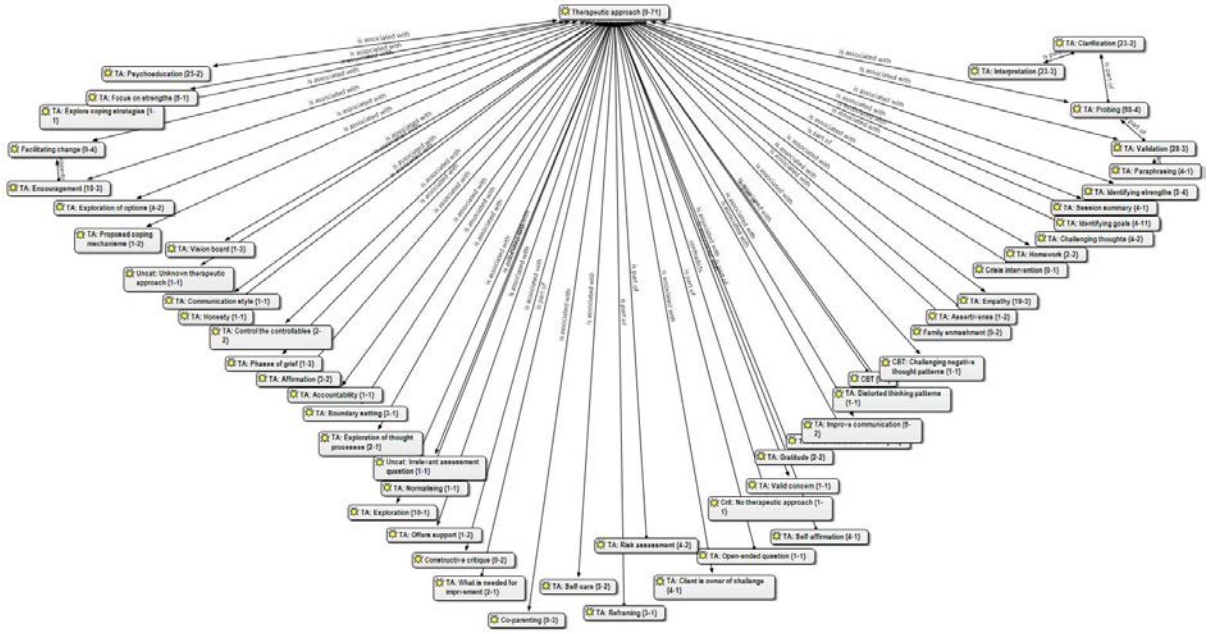


Figure 29. Therapeutic approach

The therapeutic approaches applied by the live chat agents were clustered into groups, namely i) cognitive restructuring and skill development; ii) crisis management and intervention; iii) client-

centred approach and empowerment; iv) communication and interpersonal skills; v) psychoeducation and knowledge sharing; vi) support and rapport building, and vii) problem-solving and coping strategies. The clusters with the emerging themes are displayed in Table 8.

Table 8. Therapeutic approach

Therapeutic approach (cluster)	Themes	
Cognitive Restructuring and Skill Development	Challenging negative thought patterns Challenging thoughts Distorted thinking patterns	Cognitive reframing Constructive critique Reframing
Crisis Management and Intervention	Crisis intervention	Risk assessment
Client-Centred Approach and Empowerment	Client is owner of challenge Facilitating change Empowerment Identifying goals Identifying strengths Exploration of options	Explore coping strategies Focus on strengths Self-affirmation Self-care What is needed for improvement
Communication and Interpersonal Skills	Assertiveness Empathy Encouragement Validation	Improve communication Open-ended question Paraphrasing
Psychoeducation and Knowledge Sharing	Psychoeducation Interpretation Explanation of thought processes	Homework Normalising Session summary
Support and Rapport Building	Affirmation Offers support	Gratitude Honesty
Problem-Solving and Coping Strategies	Exploration Exploration of thought processes Exploration of options	Explore coping strategies Proposed coping mechanisms Boundary setting Control the controllables

3.5.4. Theme 4: Linguistic trends

Linguistic trends of both live chat agents and the users were analysed. Results indicated the most prominent linguistic trends as Casual communication {0-2}; Clarifying questions {0-5}; Contraction Usage {0-1}; Descriptive language {0-2}; Direct questioning {0-1}; Emoji use {0-2}; Emotive language {0-1}; Emphatic {0-1}; Errors (typos/misspelling) {0-1}; Exclamation marks {0-1}; Formal language {0-2}; Friendly tone {0-1}; Impersonal {0-1}; Informative language {0-1}; Inquisitive style {0-1}; Internet slang {0-2}; Abbreviation {6-1}; Analogy {2-1}; Colloquialism {23-2}; Common expression {54-1}; Emoji {12-2}; Informal {4-9}; Laughing {1-2}; Slang {12-2}; Use of punctuation {4-1}; Vulgar language {1-1}; Linguistic trends {0-38}; Neutral tone {0-2}; Open-ended questioning {0-2}; Paraphrasing {0-1}; Personal language {0-1}; Personal pronouns {0-1}; Polite tone {0-1}; Positive language {0-1}; Reflective language {0-1}; Relaxed tone {0-3}; Respectful tone {0-2}; Rhetorical questions {0-1}; Sarcasm {0-1}; Sensitive communication {0-1}, and Specialised

knowledge/terminology {0-1}. Various examples of linguistic trends are illustrated in Excerpt 14. The relationships of the linguistic trends are displayed in Figure 30.

2(149) – Agent: Glad to see you can see the funny side of it lol :).

2(197) – User: I aim to please lol ..

2(136) - User: my dad used to get sloshed every weekend and beat up my mother in front of me.

2(235) – User: eish

2(424) – User: haibo I thought I told you it's my new supervisor and management supports him and ignore the abuse

2(18) – User: i know right! ?? I was worried

2(250) – User: Okay :-) So how will I keep in touch with you?

2(197) – Agent: Lol ..I will support the high school attitude at work until you are certain who is your friend.

2(25) – User: I'm really pissed at myself. like why would I do that?

2(440) – User: have you ever heard of this thing called a "situationshipn"

*2(362) – User: and 4 me that fact the he can just say hay f*ck u* [sic] *your mother's hole to me with no remorse and still no apology nothing before everytime he had a brink he would cry but that's changed I dnt see dat happening anymore*

Excerpt 14. Linguistic trends

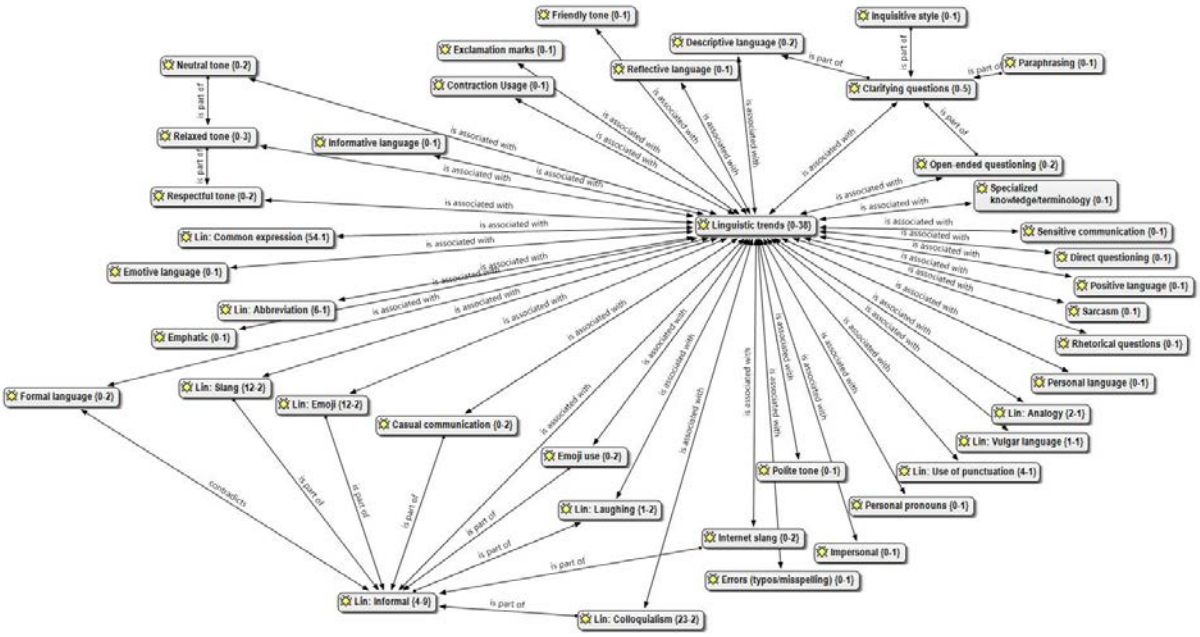


Figure 30. Linguistic trends

Linguistic styles used by both the live chat agent as well as the live chat user were clustered into groups. These include i) informal communication; ii) emotional and expressive; iii) interaction

styles; iv) written expression; v) tone and politeness; vi) creative language; vii) formal and specialised language, and viii) miscellaneous. The theme clusters with the associated themes are displayed in Table 9.

Table 9. Linguistic trends

Linguistic styles (cluster)	Themes	
Informal Communication	Casual communication	Colloquialism
	Emoji use	Informal
	Internet slang	Slang
	Abbreviation	Laughing
Emotional and Expressive	Descriptive language	Positive language
	Emotive language	Reflective language
	Emphatic	Sarcasm
Interaction Styles	Clarifying questions	Open-ended questioning
	Direct questioning	Rhetorical questions
	Inquisitive style	
Written Expression	Errors (typos/misspellings)	Exclamation marks
	Use of punctuation	Paraphrasing
Tone and Politeness	Friendly tone	Polite tone
	Impersonal	Relaxed tone
	Neutral tone	Respectful tone
Creative Language	Analogy	Emoji
	Common expression	Linguistic trends
Formal and Specialised Language	Contraction usage	Specialised knowledge/terminology
	Formal language	
Miscellaneous	Informative language	Vulgar language
	Personal language	Linguistic trends
	Personal pronouns	Specialised knowledge/terminology

From the analysed data, interactions between the live chat agent and the user seemed to follow a similar pattern - informal in nature with elements of emotive expressions. Counselling techniques, such as clarifying questions, direct questions, inquisitiveness and open-ended questions were utilised to ensure the conversational flow.

Creative language consisting of analogies, common expressions and emojis were incorporated into the chats. Formal and specialised language was occasionally used by the agents in the explanation of concepts.

3.5.5. Theme 5: Technical issues

Technical issues that, in some instances, lead to the premature conclusion of the session were identified as the end of the agent's shift. These include i) system errors; ii) loadshedding; iii)

network issues, and iv) user disconnection issues. These codes are displayed in Table 10. Where possible, the user reconnected to the session to complete the session.

Table 10. Technical

Code group	Count of code group
Agent shift ended	16
System error	1
Loadshedding	4
Network issues	4
User disconnection	64
Grand Total	**89**

'Agent shift ended' refers to an agent coming to the end of their shift, where the session was either concluded or handed over to another live chat agent to continue with the session. For example, *[2(142) – Agent: My shift is coming to an end, My colleague will read the chats and take over if it is okay with you].*

'System errors' refers to any error that the live chat platform could experience and that could inhibit the completion of the live chat session. For example, *[2(444) – Agent: Our system is down so I can't check to see if have an existing profile for you].*

'Loadshedding' either results in network issues or in the user's mobile phone battery being depleted, hindering them from completing the session. For example, *[2(100) – User: there' s* [sic] *loadshedding I have network issue].*

'Network issues' refer to any network provider that the user or agent might experience that might prohibit them from completing the session. For example, *[2(171) – Agent: We do seem to have an unstable connection –the chats are disconnecting].*

'User disconnection' refers to circumstances in which the client's mobile phone disconnects from the live chat platform. User disconnections occur under numerous conditions, such as when the user moves out of a Wi-Fi reception area or when the user's phone disconnects. It has been found that, more often than not, the user reconnects to complete the session. The following is a standardised notification from several sessions:

> *[Please kindly note that with regards to chat connectivity, please note that the chat will disconnect if you take a phone call/close the chat window, lock your phone screen, or move out of data reception. The new chat then loses the previous history. For best results we require continuous interaction. Please also ensure your App has been updated to the latest version.]*

3.6. SENTIMENT ANALYSIS: LIVE CHAT TRANSCRIPT ANALYSIS

NLP research in the field of sentiment analysis focuses on determining the attitude, opinion or mood of human subjective aspects in texts (Bhadane *et al.*, 2015; Pandian, 2021). The sentiment analysis in this research was conducted using ATLAS.ti AI, employing techniques grounded in ML and NLP (Yadav & Vishwakarma, 2020).

Table 11 and Figure 31 show the analysis of sentiments extracted from live chat counselling transcripts, consisting of both agent and user texts. The findings revealed that among the sentiments examined, a substantial portion (47%) emerged as negative, shedding light on the challenges and concerns expressed by participants. Notably, 25% of the sentiments maintained a neutral tone, capturing instances where emotions were less pronounced or information-driven. Additionally, a noteworthy 28% of the sentiments conveyed positivity, reflecting the moments of encouragement, progress, and optimism experienced during the counselling interactions.

Table 11. Sentiment analysis

Sentiment	Count of sentiment
Sentiment: Negative	3030
Sentiment: Neutral	1639
Sentiment: Positive	1809
Grand Total	**6478**

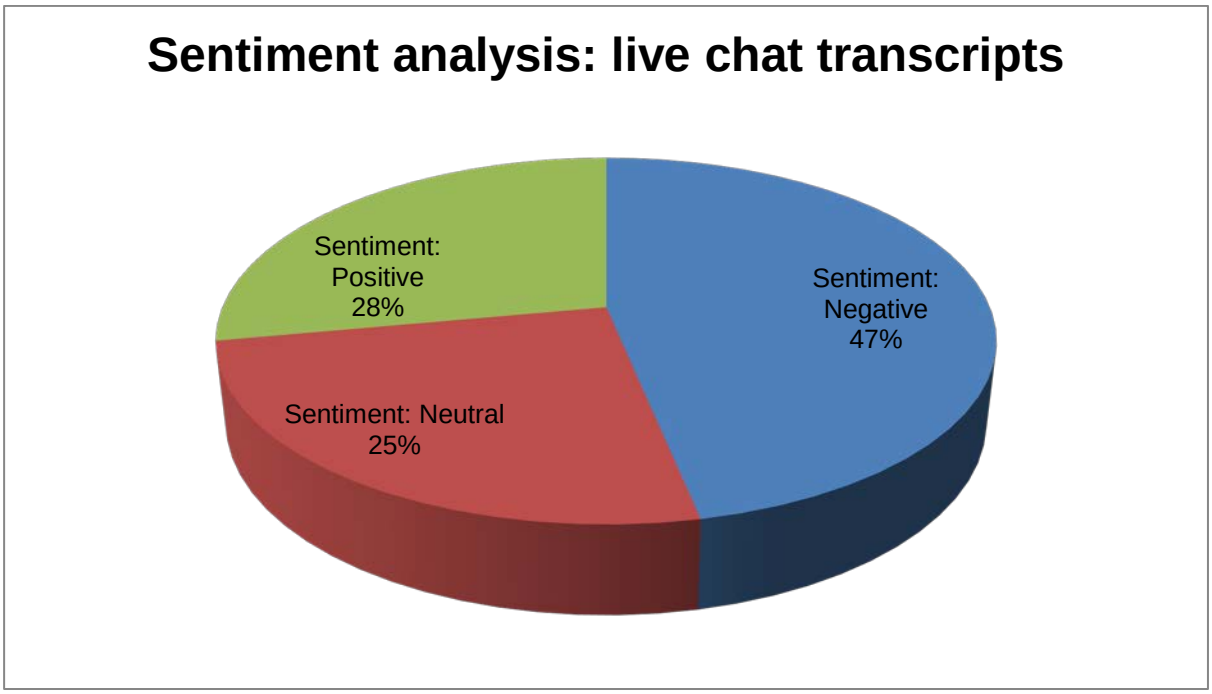


Figure 31. Sentiment analysis

3.6.1. Negative sentiments

According to Gandhi *et al.* (2022), negative sentiments refer to the expression of unfavourable emotions, attitudes, opinions or feelings within the analysed text. These sentiments indicated that the user or agent conveyed dissatisfaction, distress or pessimism regarding their challenges, experiences or concerns. The sentiments often reflected the presence of emotional distress, dissatisfaction or critical feedback during the live chat counselling sessions. In this analysis, 47% of sentiments were identified as being negative, indicating the prevalence of negative emotions and concerns amongst users and live chat agents. Excerpt 15 illustrates negative sentiments.

2(5) – User: I'm not sure where or even how to begin I just feel lost and confused about my life I just can't cope

2(82) – User: it's thrown me off everything completely, it's all kind of overwhelming. It's been about a month since I got the court order and the children are rumored to be leaving this month. Little things bring me to my knees at times and I've been taking those moments as they come. But it hasn' t gotten any easier I don' t know if I'm managing.

(2)136 – User: when Iused to get too upset with my toddler I shout at him. and he just cries. that doesn't help. so I started explaining to him when he does wrong but he doesn't pay attention to me. and then it frustrates me but Iwalk away from him to calm down and decide what to do. but get emotional thinking i cant handle to situation

2(228) – User: another bcz we both went thru some stuff like she knew my husband was beating me (even whilst I was pregnant I had to go to hospital) etc and then when I was diagnosed with cancer he was not supportive at all. He never took me to chemo, never spent time with me ...had fights with me blah blah blah ....

2(294) – User: we are married for 9 years we have always been having problems with him drinking and behaving so badly but within the last year the drinking has gone very bad. he drinks every chance he gets and he doesn't stop. can drink going through to the next day and then becomes violent

Excerpt 15. Negative sentiments

3.6.2. Neutral sentiments

Neutral sentiments, as identified in the sentiment analysis, represented instances where the language used in live chat counselling sessions was relatively free of strong emotional expressions (Gandhi *et al.*, 2022). These sentiments captured moments when users or agents conveyed information or opinions in a relatively neutral or objective tone of voice, with emotions being less pronounced. Essentially, they indicated the absence of strong positive or negative feelings. The analysis found that 25% of the utterances in the transcripts suited this category, indicating cases where the counselling interactions were primarily focused on the exchange of information or where emotions did not come to the fore. Excerpt 16 illustrates neutral sentiments.

2(40) – Agent: How have you been reflecting exactly? There is no right or wrong way to reflect. Everyone is different and deals with their emotions differently.

2(82) – User: we were not in a relationship. But I filed for a protection order after the children' s mother got physical for a second time in my mothers house.

2(87) – User: I have had discussions with my line manager to explain the diagnosis and impact without divulging too much information, as I believe there needs to be a separation of work and personal life.

2(197) – User: I only have 1 person who I think is a good friend for work which is my female coworker .. we started at the same time so, when we were talking we shared the same opinion about the work environment.. I don' t necessarily trust her, but she makes the place a bit bearable if she wasn' t there to help me I would' ve left yesterday already but I can't because I have responsibilities.

2(262) – Agent: Rushing into things will create more difficulty. This is a brand new person that has nothing to do with the previous partner.

Excerpt 16. Neutral sentiments

3.6.3. Positive sentiments

Positive sentiments within the context of this analysis represented the expression of favourable emotions, attitudes and opinions (Gandhi *et al.*, 2022) within the analysed live chat counselling transcripts. These sentiments represented instances where participants expressed satisfaction, encouragement, progress or optimism concerning their counselling experiences or private circumstances. The analysis found that 28% of the sentiments matched this category, highlighting the occurrence of uplifting feelings and energising interactions during counselling sessions. Excerpt 17 illustrates positive sentiments.

2(5) – User: I will thank you so much really I feel so much better really Agent It is a great pleasure. do feel free to reach out at anytime. User I will thank you

2(19) – Agent: I encourage you to please practice self-care. It is easy to get so trapped in the cycle of caring and supporting someone that you forget about yourself.

2(437) – User: Mostly my support structure and for a long time that consisted of my family, friends and scholarship

2(447) – User: Must I tell him how I feel and be as honest like I was with you now?

2(504) - User and he is 13years my senior and he has 2 kids whom i considered my own.

Excerpt 17. Positive sentiments

3.7. TEXTUAL TRENDS

Live chat transcripts were subjected to a word crunch query. The query resulted in 6,336 words being analysed with a cumulative occurrence of 247,268. The words included in the results had a cumulative total of 20,126 occurrences. This query provided each word and its occurrence throughout the data set. Stop words were removed and the most prominent words (frequency >0.01%) that have significant meaning to mental health were included in the results, as displayed in Table 12.

Table 12. Textual trends

Cluster and cluster frequency	Word and word frequency	
Emotions and Feelings (Total: 5.96%)	feel (0.58%) feeling (0.17%) emotions (0.09%) feelings (0.09%) happy (0.09%) sad (0.02%) anxious (0.03%) overwhelmed (0.03%) anger (0.03%) guilt (0.01%) numb (0.01%)	suicidal (0.01%) suicide (0.01%) panic (0.01%) stress (0.07%) depressed (0.02%) hate (0.01%) fear (0.01%) hope (0.01%) hurt (0.04%) love (0.10%) proud (0.02%)
Relationships and Family (Total: 3.57%)	relationship (0.18%) care (0.11%) family (0.11%) mom (0.06%) dad (0.05%) mother (0.05%) father (0.04%) partner (0.04%) husband (0.07%) wife (0.03%) friend (0.04%) friends (0.06%) boyfriend (0.02%) girlfriend (0.01%) child (0.04%)	children (0.03%) son (0.03%) daughter (0.04%) marriage (0.01%) divorce (0.02%) couples (0.01%) cheating (0.02%) respect (0.02%) aunt (0.01%) cousin (0.01%) bestie (0.01%) sisters (0.01%) sexual (0.01%) brothers (0.01%) siblings (0.01%)
Self and Personal Growth (Total: 0.55%)	self (0.05%) yourself (0.32%) believe (0.06%) control (0.06%) change (0.09%) growth (0.01%) goals (0.02%)	strengths (0.01%) personality (0.01%) strength (0.01%) lessons (0.01%) forgiveness (0.01%) meditation (0.02%)
Mental Health and Wellbeing (Total: 1.60%)	mental health (0.04%) depression (0.02%) bipolar (0.01%) anxiety (0.06%) stress (0.07%) trauma (0.01%) addiction (0.01%) alcohol (0.01%) smoking (0.01%)	medication (0.01%) meds (0.01%) therapy (0.07%) counselling (0.13%) help (0.19%) support (0.13%) self-care (0.03%) self-help (0.01%) recovery (0.01%)
Work and Career (Total: 0.82%)	work (0.38%) job (0.01%) working (0.09%) career (0.03%) income (0.01%) financial (0.05%)	salary (0.02%) employment (0.01%) boss (0.01%) education (0.01%) exams (0.01%) workplace (0.01%)
Life Events and Challenges (Total: 0.36%)	time (0.44%) life (0.17%) future (0.02%) past (0.02%)	passing (0.01%) grief (0.03%) change (0.09%) loss (0.02%)

	decisions (0.01%) challenges (0.02%) difficulties (0.01%) struggles (0.01%) adversity (0.01%) accident (0.01%)	trauma (0.01%) debt (0.02%) legal (0.02%) court (0.02%) hospital (0.01%)
Personal Habits and Lifestyle (Total: 0.10%)	habits (0.01%) addiction (0.01%) smoking (0.01%) drinking (0.01%) diet (0.01%) exercise (0.01%)	sleep (0.01%) stress (0.07%) healthy (0.03%) unhealthy (0.01%) self-medicating (0.01%)
Social and Environmental Factors (Total: 0.27%)	social (0.02%) environment (0.02%) society (0.01%) culture (0.01%)	gender (0.02%) friends (0.01%) school (0.01%) workplace (0.01%)
Support and Coping Strategies (Total: 1.08%):	support (0.13%) coping (0.02%) coping strategies (0.01%) journaling (0.02%) meditation (0.02%) self-help (0.01%) self-care (0.03%) self-medicating (0.01%)	positive (0.04%) negative (0.03%) overwhelm (0.03%) stressors (0.03%) pressure (0.03%) challenge (0.03%) adapt (0.01%)

3.8. DISCUSSION OF RESULTS AND DATA CONVERGENCE

Inferences were made from the analysed data, namely that live chat sessions followed a near-identical pattern. First, the live chat user initiated the session. Thereafter, the agent replied with a response that correlated to whether the user was a new user or a returning user. These therapeutic sessions also followed a pattern of an initial phase; a working phase, and an ending phase.

Data were gathered from all phases, with rich data emerging from the working phases. The most prominent themes that emerged were i) problems/challenges; ii) counsellor approaches to the sessions; iii) linguistic trends; iv) technical aspects; v) sentiments expressed and vi) textual trends, as expressed in Figure 32.

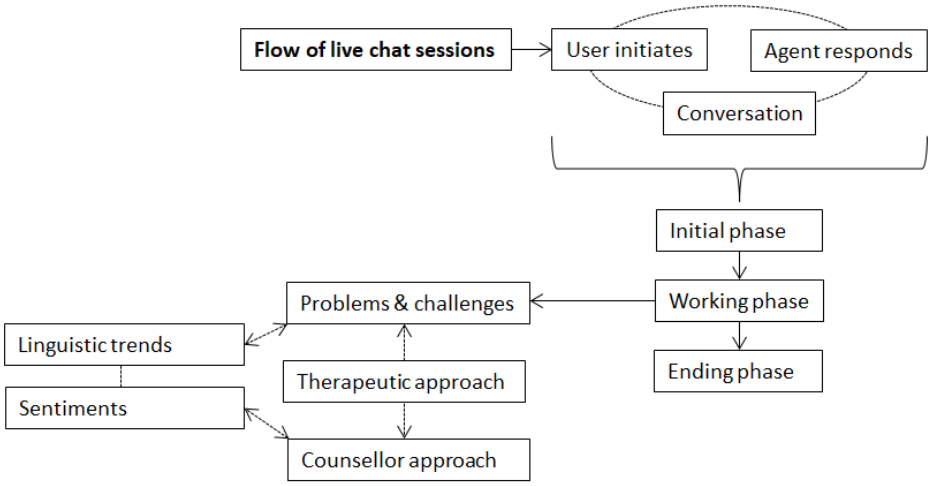


Figure 32. Live chat session flow and elements

Problem/challenge statements (1,610) were identified as career and work-related issues, dysfunction, emotional wellbeing, family dynamics, financial issues, interpersonal relationships, personal development, psychological challenges, and requests for help/support. Problem discussion was the focal point of interactions between the user and agent. This was strongly resonated by the sentiment analysis, which indicated that 3,030 sentiments that were expressed were negative, with 1,639 being neutral. Textual trends furthermore indicated a focus on problems/challenges with problem-related linguistic use being emotions and feelings, relationships and family, self and personal growth, mental health and wellbeing, work and career, life events and challenges, personal habits and lifestyle, social and environmental factors, and support and coping strategies. Linguistic styles employed by both user and agent to convey messages included informal communication, emotional expressions, questions, written expressions, different tones and degrees of politeness, formal and specialised language and miscellaneous trends.

Agents employed counselling approaches relating to advice and support (63) and psychological approaches (39) to assist users. Therapeutic approaches were also employed. They consisted of cognitive restructuring and skills development, crisis management/intervention, client-centred approach, psychoeducation and knowledge sharing, support and rapport building, and problem-solving and coping strategies.

Word frequencies indicated an inclination towards the use of words that predominantly related to emotions and feelings, relationships, family, self- and personal growth, mental health, wellbeing, work and career, life events and life challenges, personal habits, lifestyle, social and environmental factors, and support and coping strategies. The word usage was distributed among different sentiment categories, namely 28% in positive sentiments, 47% in negative sentiments and 25% in neutral sentiments.

3.9. RECOMMENDATIONS

Based on the extensive analysis of live chat counselling transcripts, it is recommended that consideration be given to the development of an AI generative counselling chatbot framework. The findings confirmed a need for accessible and effective mental health support. The structure observed in the analysis of live chart counselling transcripts can be effectively harnessed and enhanced through the implementation of AI.

The analysis identified an array of negative sentiments, which are indicative of emotional distress and numerous challenges faced by live chat counselling users. The demand for mental healthcare support continues to increase, indicating that AI chatbots could be a solution to offer immediate assistance and support.

The structured three-phase model (observed in the live chat transcript analysis), consisted of an initial-, working- and ending phase. The findings aligned with the capabilities of AI chatbots to engage in structured conversations. The synchrony could, moreover, ensure a positive user support interface.

The potential to offer mental health help to a larger target group is made possible by the scalability and cost-effectiveness of AI chatbots. It is possible to accommodate multiple users simultaneously, thereby guaranteeing that more services are provided to those that require them. AI chatbots are the ideal response as they provide instant support, allowing users to express their issues and receive help outside of typical business hours. This is especially helpful in times of crisis, as timely intervention is crucial for mental health support. In addition, anonymity and privacy are highly valued by users, providing the opportunity to discuss sensitive issues and topics.

This study does not insinuate or recommend that human counsellors be replaced. Rather, it proposes that AI counselling chatbots can fulfil a supplementary role.

An AI framework can generate valuable insights from the vast amount of data it collects, which can contribute to the continuous enhancement of mental health services. The development of an AI counselling framework opens up opportunities for research and development in the field of mental health and technology and may lead to innovative solutions to address mental health concerns.

The development of an AI generative counselling framework possesses the potential to enhance the accessibility and quality of mental healthcare services. A proactive response to the rising demand for mental healthcare services is potentially offered with an emphasis on delivering timely, confidential and personalised assistance to mental healthcare users.

3.10. CONCLUSION

The empirical findings indicated that live chat counselling sessions typically commenced with the user initiating the conversation, whereafter the counselling session proceeded through three phases (initial, working and ending) in a structured fashion. The initial phase usually consisted of introductions and a problem statement. The working phase or the main counselling phase involved a two-way conversation between the counsellor and the agent, whereby support strategies and advice were offered. The ending or conclusion phase consisted of either handing over to another counsellor for further support or ensuring there was a clear conclusion to the session.

The structured approach of the live chat sessions ensured effective two-way communication between the counsellor and the user, allowing a flowing exchange of information and problem-solving within the confines of a text-based live chat platform. Additionally, the analysis revealed several noteworthy themes that provided insight into the subjects and topics that were discussed during live chat sessions. Users tended to discuss various challenges and concerns during sessions, spanning career- and work-related issues, emotional wellbeing, family dynamics, financial concerns, interpersonal relationships, personal development, psychological challenges, and requests for help and support.

Counsellors drew upon various counselling approaches and techniques to address user concerns. The approaches include offering advice and support, using psychological methods to address emotional- and mental health concerns and the use of therapeutic techniques to guide users in addressing concerns and challenges.

Linguistic trends indicated elements relating to informal communication, emotional expressions, questioning, written expressions, tone differences, various degrees of politeness and formal-specialised language usage. Furthermore, a sentiment analysis was conducted on counsellor- and user sentiments. A high prevalence, suggestive of emotional distress, was found. This is consistent with the notion that the live chat platform is utilised to express distress and emotional difficulty.

A word frequency analysis was conducted and revealed the most frequently used words in live chat counselling sessions. The frequency analysis related to a variety of issues commonly discussed in live chat counselling sessions. Words were distributed along various categories, including emotions and feelings, relationships, personal growth, mental health, work and career, life events and life challenges, personal habits and lifestyle, social- and environmental factors, and support and coping strategies.

REFERENCES

Andersson G. 2016. Internet-delivered psychological treatments. *Annu Rev Clin Psychol*, 12:157–79.

Argyropoulou, K. & Kaliris, A. 2018. From career decision-making to career decision-management: New trends and prospects for career counseling. *Advances in Social Sciences Research Journal*, 5(10):483-502.

Ashton, K. 2009. That 'internet of things' thing. *RFID journal*, 22(7):97-114.

Augar, N. & Zeleznikow, J. 2014. Developing online support and counseling to enhance family dispute resolution in Australia. *Group Decision and Negotiation*, 23:515-532.

Baldwin, R.E. 2019. *The Globotics Upheaval: Globalization, Robotics, and the Future of Work*. New York, NY: Oxford University Press.

Bauer, J. 2021. *The Transformative Self: Personal Growth, Narrative Identity, and the Good Life*. New York, NY: Oxford University Press.

Bhadane, C., Dalal, H. & Doshi, H., 2015. Sentiment analysis: Measuring opinions. *Procedia Computer Science*, 45:808-814.

Binder, P., Moltu, C., Hummelsund, D., Sagen, S.H. & Holgersen, H. 2011. Meeting an adult ally on the way out into the world: Adolescent patients' experiences of useful psychotherapeutic ways of working at an age when independence really matters. *Psychotherapy Research*, 21(5):554–566.

Cowls, J., Tsamados, A., Taddeo, M. & Floridi, L. 2021. A definition, benchmark and database of AI for social good initiatives. *Nat Mach Intell*, 3:111–115.

Doyle, L., McCabe, C., Keogh, B., Brady, A. & McCann, M. 2019. An overview of the qualitative descriptive design within nursing research. *Journal of Research in Nursing*. 25(5):443-455.

Fava, N.M., Schuler, B.R. & Kelly, M.S. 2016. The incorporation of big data in social work practice: Current efforts and future directions. *Journal of Social Work Education*, 52(3):316-328.

Fernández, E.N. & Mairal, J.B. 2017. Behavioral activation versus cognitive restructuring to reduce automatic negative thoughts in anxiety generating situations. *Psicothema*, 29(2):172-177.

Fiske, A., Henningsen, P. and Buyx, A. 2019. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of medical Internet research*, 21(5), e13216. https://www.jmir.org/2019/5/e13216/

Fonseka, T.M., Bhat, V. & Kennedy, S.H. 2019. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Australian & New Zealand Journal of Psychiatry*, 53(10):954-964.

Foster, I., Ghani, R., Jarmin, R.S., Kreuter, F. & Lane, J. 2020. *Big Data and social science: Data science methods and tools for research and practice*. CRC Press.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L. & Rauws, M. 2018. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health*, 5(4), e9782. https://mental.jmir.org/2018/4/e64/

Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J. & Sartorius, N. 2015. Toward a new definition of mental health. *World psychiatry*, 14(2):231-233.

Gandhi, A., Adhvaryu, K., Poria, S., Cambria, E. & Hussain, A., 2022. Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion*, 91:424-444.

Gibson, K. & Cartwright, C., 2014. Young people's experiences of mobile phone text counselling: Balancing connection and control. *Children and youth services review*, 43:96-104.

Gibson, K. & Cartwright, C. 2013. Agency in young clients' narratives of counseling: 'It's whatever you want to make of it'. *Journal of Counseling Psychology*, 60(3):340–352.

Giglio, F. 2021. Fintech: A literature review. *European Research Studies Journal*, 24(2B):600-627.

Goldkind, L. 2021. Social Work and Artificial Intelligence: Into the Matrix. *Social Work,* 66(4):372–374.

Güzel, B. 2021. Futurism in Social Work: Transhumanism, Gamification of Human Life, Homo Roboticus and Care. *Journal of Society & Social Work*, 32(2):589-606.

Hajjaji, Y., Boulila, W., Farah, I.R., Romdhani, I. & Hussain, A. 2021. Big data and IoT-based applications in smart environments: A systematic review. *Computer Science Review*, 39:1-39.

Hull, T.D., Malgaroli, M., Connolly, P.S., Feuerstein, S. & Simon, N.M. 2020. Two-way messaging therapy for depression and anxiety: longitudinal response trajectories. *BMC psychiatry*, 20(1):1-12.

Inkster, B., Sarda, S. & Subramanian, V. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR Mhealth Uhealth*, 6(11):3-18.

Javaid, M., Haleem, A., Singh, R.P. & Suman, R. 2022. Artificial intelligence applications for industry 4.0: A literature-based study. *Journal of Industrial Integration and Management*, 7(1):83-111.

Koutsouleris, N., Hauser, T.U., Skvortsova, V. and De Choudhury, M. 2022. From promise to practice: towards the realisation of AI-informed mental healthcare. *The Lancet Digital Health*, 4(11):829-840.

Kupferberg, A., Bicks, L. & Hasler, G. 2016. Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, 69:313-332.

Lee, E.E., Torous, J., De Choudhury, M., Depp, C.A., Graham, S.A., Kim, H.C., Paulus, M.P., Krystal, J.H. & Jeste, D.V. 2021. Artificial intelligence for mental healthcare: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(9):856-864.

Lim, S.M., Shiau, C.W.C., Cheng, L.J. & Lau, Y. 2022. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. *Behavior Therapy*, 53(2):334-347.

Lyons, S.T., Schweitzer, L. & Ng, E.S. 2015. Resilience in the modern career. *Career Development International*, 20(4):363-383.

Magogodi, P.P.S. 2017. *The Use of Online Text Based Technologies as a Medium for Employee Counselling: Perceptions of Online Counsellors*. Johannesburg: University of the Witwatersrand. (Dissertation – MA).

Marcelle, E.T., Nolting, L., Hinshaw, S.P. & Aguilera, A, 2019. Effectiveness of a multimodal digital psychotherapy platform for adult depression: a naturalistic feasibility study. *JMIR mHealth and uHealth*, 7(1), e10948. https://mhealth.jmir.org/2019/1/e10948

Marsch, L.A. 2021. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacol.* 46:191–196.

Mathekga, H.L. & Sekudu, J. 2019. Employee wellness: Preparation for the fourth industrial revolution. *International Journal of Innovative Studies in Sociology and Humanities*, 4(5):31-39.

Mathieu, S.L., Uddin, R., Brady, M., Batchelor, S., Ross, V., Spence, S.H., Watling, D. & Kõlves, K. 2021. Systematic review: The state of research into youth helplines. *Journal of the American Academy of Child & Adolescent Psychiatry*, 60(10):1190-1233.

Michel, N.M., Rowa, K., Young, L. & McCabe, R.E. 2016. Emotional distress tolerance across anxiety disorders. *Journal of anxiety disorders*, 40:94-103.

Morrissey, D. 2021. *Social Work and Helping Humanity Through Chatbots*. https://medium.com/@devinmorrissey.1987/social-work-and-helping-humanity-through-chatbots-a669dab63b4d Date of access: 14 Jan. 2023.

Ndzendze, B., Singh, A., Timm, S., Meyers, C., Malatji, M.J., Malatji, M.J., … Brokensha, S.I. 2024. *The 4IR and the Humanities in South Africa: Perspectives on innovation, power and potentialities*. Cape Town: AVARSITY Books.

Pandian, A.P. 2021. Performance evaluation and comparison using deep learning techniques in sentiment analysis. *Journal of Soft Computing Paradigm*, 3(2):123-134.

Patel, K.K., Patel, S.M. & Scholar, P. 2016. Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. *International journal of engineering science and computing*, 6(5):6122-6131.

Rice, E., Wilder, B., Onasch-Vera, L., DiGuiseppi, G., Petering, R., Hill, C., Yadav, A., Lee, S.J. & Tambe, M. 2021. A Peer-Led, Artificial Intelligence–Augmented Social Network Intervention to Prevent HIV Among Youth Experiencing Homelessness. *Journal of Acquired Immune Deficiency Syndromes*, 88(1):20-26.

Roestenburg, W.J.H. 2021. *Analysing technology acceptance by social workers in South Africa*. Potchefstroom: North-West University. (Dissertation – M.B.A.).

Safodien, M. 2021. Social work 4.0? The fourth industrial revolution and social work education: a South African perspective. *Social Work/Maatskaplike Werk*, 57(3):257-271.

Shreyas, J., Jumnal, A., Kumar, S.D. & Venugopal, K.R. 2020. Application of computational intelligence techniques for internet of things: an extensive survey. *International Journal of Computational Intelligence Studies*, 9(3):234-288.

Slomski, A., 2015. Online therapy reduces suicide ideation in medical interns. *Jama*, 314(24):2608-2618.

Sunter, C. & Ilbury, M. 2021. *Thinking the future: New Perspectives from the shoulders of giants.* Cape Town: Penguin Books.

Tambe, M. & Rice, E. 2018. Merging Social Work Science and Computer Science for Social Good. In: Tambe, M & Rice, E., eds. *Artificial Intelligence and Social Work*. Cambridge: Sheridan Books. pp. 48-72.

Srinivasan, J. 2021. *2 Ways AI is Transforming Wellness in the Workplace*. https://www.spiceworks.com/hr/engagement-retention/guest-article/2-ways-ai-is-transforming-wellness-in-the-workplace/ Date of access: 14 Jan. 2023.

Topooco, N., Byléhn, S., Dahlström, E., Holmlund, J., Lindegaard, J… Andersson, G. 2019. Evaluating the efficacy of internet-delivered cognitive behavioral therapy blended with synchronous chat sessions to treat adolescent depression: randomized controlled trial. *Journal of medical Internet research*, 21(11), e13393. https://www.jmir.org/2019/11/e13393

Topooco, N., Berg, M., Johansson, S., Liljethörn, L., Radvogin, E., Vlaescu, G… Andersson, G. 2018. Chat-and internet-based cognitive–behavioural therapy in treatment of adolescent depression: randomised controlled trial. *BJPsych open*, 4(4):199-207.

Tutun, S., Johnson, M.E., Ahmed, A., Albizri, A., Irgil, S., Yesilkaya, I,… 2023. An AI-based decision support system for predicting mental health disorders. *Information Systems Frontiers*, 25(3):1261-1276.

Toyama, M., Fuller, H.R. & Hektner, J.M. 2020. Psychosocial Factors Promoting Personal Growth Throughout Adulthood. *Journal of Happiness Studies,* 21:1749-1769.

Yadav, A. & Vishwakarma, D.K., 2020. Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, 53(6):4335-4385.

CHAPTER 4: PROPOSED PRODUCT: AN INTELLIGENT ONLINE COUNSELLING FRAMEWORK FOR AN EMPLOYEE HEALTH AND WELLNESS PROGRAMME

4.1. INTRODUCTION

The convergence of the findings of the literature study (Chapter 2) and empirical study (Chapter 3) found that for social work to remain relevant as a profession and academic field, it needs to adopt digital technology to join the worldwide technology ecosystem. Social work practitioners must embrace data management tools, implement digital interventions, develop digital competencies, and establish collaboration between mental healthcare professionals and IT- and HIT experts. To advance the field, practitioners must prioritise ongoing research and adopt user-centred design principles in developing future digital interventions.

In this chapter, an intelligent online counselling framework designed to enhance EHWPs is proposed. The chapter builds on the findings from previous chapters and details goals, technical requirements, and user engagement strategies that are necessary for the framework's implementation. Findings from the excerpts will be used to underline the processes required in developing the framework. These excerpts are shown verbatim, with erroneous grammar and spelling that highlight both the users' and agents' cadences when using the chatbot.

The empirical findings outlined in Chapter 3 underscore the necessity to develop an intelligent online counselling framework. This framework aims to inform an AI mental health chatbot to meet the specific needs of users within the EHWP. By fostering comprehensive employee health, such a framework possesses the potential to improve the effectiveness and accessibility of counselling services to meet the demands of rapidly modernising workplaces.

To design this framework, the SDLC (Royce, 1970) - which has been confirmed as the most appropriate approach - was followed. The term SDLC refers to a process for organising, designing, coding, testing, and implementing user requirement specifications and is regarded as a key concept in software engineering. The SDLC encompasses various hardware and software configurations based on a methodical approach to producing high-calibre software for end users. Planning, analysis, design, coding, testing, implementation and evaluation are all the necessary sequential steps for software engineers to complete a configuration (Sharma, 2017).

Writing effective code for the creation of high-quality and functional programs is merely one of the processes that software engineering entails. Thus, a well-defined SDLC model facilitates the creation of high-quality software (Akinsola *et al.*, 2020). The SDLC has been described in various ways, with the descriptions varying depending on which field a system has been created for. However, for all fields, a quality, useable product is developed using the life cycle approach. The strategy is a flexible development method, allowing managers, designers, and clients to engage directly in the creation of the final product, unlike more rigid, sequential processes that

are less user-driven (Pinciroli *et al.* 2022). The primary SDLC model historically employed in software engineering endeavours to guarantee project efficacy is commonly known as the waterfall model. Within the framework of the waterfall model, the entirety of the software development process is delineated into discrete phases. Typically, within this sequential progression, the output of one phase serves as the precursor input for the subsequent phase.

The SDLC waterfall model typically consists of six to eight phases (Christanto & Singgalen, 2023; Saravanos & Curinga, 2023), which was not fully completed for this study. The noted phases consist of a i) planning phase; ii) analysis phase; iii) design phase; iv) development phase; v) testing phase; vi) maintenance and support phase, and vii) the evaluation phase. For this study, only three phases (planning, analysis and a partial design) were completed. Further phases would have necessitated the collaboration of specialists in the field of IT and HIT. This collaborative approach superseded the confinement of this study.

To refine this developed AI generative framework, the Delphi technique was implemented. The Delphi technique facilitated a systematic consultation process with subject matter experts to achieve consensus on the framework's design and functionality.

4.2. DELPHI PANEL METHODOLOGY FOR REFINING THE AI COUNSELLING FRAMEWORK

The final phase of the study focused on the refinement of the AI counselling framework through the implementation of the Delphi technique. This technique involved the engagement of subject matter experts in iterative rounds of feedback to improve content validity. Psychosocial health and ICT experts were included in the study to ensure relevant and diverse expertise (Kasi *et al.*, 2008).

A draft of the proposed framework was systematically refined by conducting four rounds of the Delphi process. In the first round, participants provided the initial input and contributed additional pertinent information. The second round aimed to achieve consensus on understanding the proposed framework and included recommendations for improvement. The third examined disagreements, agreements and the underlying premises that informed different perspectives—the final round involved evaluating and adapting the framework, followed by summative feedback to the participants. To optimise efficiencies and minimise logistical complexities, the sessions were conducted via Microsoft Teams, leveraging the e-Delphi technique (Brüggen & Willems, 2009). Combining discussion groups and Delphi methodology, enhanced the reliability of collected data while mitigating potential bias (Mead and Mosely, 2001). Consensus was reached after four rounds through systematic feedback and iterative adjustments to the framework, guided by expert input.

Experts were selected through purposive sampling, focusing on professionals in psychosocial health and ICT with credible qualifications, experience and practical work within their respective

fields. Recruitment was conducted by an independent consultant functioning within Lyra's affiliate network to avoid conflicts of interest. Inclusion criteria required participants to be affiliated with Lyra SA (as external affiliates or practitioners), possess relevant qualifications and practical experience in psychosocial health, ICT, or digital technologies. The expert panel was purposefully limited to ten members to diversify input while limiting the potential risk of participants influencing each other's perspectives.

Participants were recruited through the distribution of advertisements distributed by Lyra SA's clinical manager, who acted as mediator. Written informed consent was obtained, and confidentiality was ensured by anonymising participants through the removal of all identifiable information.

Ten Delphi panellists participated in the refinement of the proposed AI counselling framework. An initial draft of the proposed framework was distributed to the panellists via email. Panellists were requested to submit written feedback on the content and design. The subsequent Delphi panel group discussions were conducted via Microsoft Teams over a period of four weeks. The initial session comprised of a summative presentation of the proposed product to ensure that all panellists shared a fundamental understanding. Sessions spanning from the second to the fourth rounds involved unstructured group discussions, which allowed for open dialogue and iterative refinement. To extract meaningful insights, all sessions were transcribed and analysed. The analysed outcomes of the group discussions were systematically integrated into the proposed product. Consensus was reached during the final round, after which the proposed product was refined according to the recommendations. As conclusion, a summative presentation of the finalised product was shared with the Delphi panel. The characteristics of the Delphi panellists, including their areas of expertise, professional affiliations, contribution focuses, and attendance across the respective sessions or rounds, are presented in Table 13.

Table 13. Delphi panellist expertise, affiliations, and key contributions

Panellist	Specialisation	Professional affiliation	Key contributions
01	Social worker	Lyra affiliate	• Empathy • Human connection • Lack of personal connection • Maintaining human element
02	Social worker	Lyra affiliate	• Risk assessment • Referral mechanism
03	Registered counsellor	Lyra affiliate	• Psychosocial • Support to traditional service
04	Social worker	Lyra affiliate	• Ethical • Cultural sensitivity • Language • Adverse event reporting
05	Social worker	Lyra affiliate	• Clinical application • Depth of clinical relationship

			• Human witnessing
06	Social worker	Lyra affiliate	• Psychoeducation • Skills development
07	Registered counsellor	Lyra affiliate	• Clinical theory • Limitations in probing • Lack of intuition
08	Social worker	Lyra affiliate	• Accessibility
09	Registered counsellor	Lyra affiliate	• Truthfulness • Response to emotional cues
10	Application developer	Lyra consultant	• Practicality of implementation • Costing • Data costs • Benefits • Limitations • Model training • Application integration • Interface considerations

Data analysis employed qualitative methods, with experts providing detailed inputs on the framework design. The finalisation of the Delphi process was signalled by reaching consensus amongst expert inputs. The iterative feedback loop enabled the refinement and validation of the proposed framework, ensuring that the proposed product conformed to the standards and expectations of the expert panel.

Reliability, validity and trustworthiness were ensured through the structured approach to consensus building offered by the Delphi technique. Refinement and improvement of the framework were ensured by inputs obtained through multiple rounds.

Through the structured and iterative properties of the Delphi technique, reliability, validity and trustworthiness were ensured. Expert inputs fostered the rigorous refinement and alignment of multidisciplinary standards, facilitated by the phased approach. The practical applicability, and relevance of the framework within the field of social work and EHWPs were enforced by the validation of the framework's content through consensus achieved by the experts. The robust process and methodological approach ensured that the proposed framework meets both theoretical and practical expectations in its intended application. The results of the Delphi discussion groups are presented in the various sections of the framework discussion. The following section presents the refined intelligent online counselling framework that emerged from the Delphi discussion group.

4.3. THE REFINED INTELLIGENT ONLINE COUNSELLING FRAMEWORK

The development of the intelligent online counselling framework consists of a planning phase, an analysis phase, and clinical theoretical requirements.

The planning phase defines the framework goals, which include enhancing accessibility, addressing user-specific challenges, and employing evidence-based counselling approaches that

grounded in sound psychological theories and linguistic patterns. In the analysis phase, problems and functional needs are identified through the incorporation of thematic analysis, with findings integrated to inform the framework requirements. The clinical theoretical requirements encompass the integration of advanced AI features, dynamic risk assessment, theoretical approaches, referral mechanisms, cultural sensitivity, evidence-based assessments, Bayes' theorem, ethical integrity, and external system integration.

4.3.1. Planning phase

The initial phase of the Software Development Life Cycle (SDLC) involves setting clear goals and objectives as a fundamental step (Christanto & Singgalen, 2023; Saravanos & Curinga, 2023). The goal of the initial phase (planning phase) was to define the goals and objectives of an intelligent online counselling framework for an EHWP, based on findings and analysis of the empirical data. In total, seven overarching goals for an intelligent online framework were identified. The identified goals include developing a system that i) offers accessible counselling to everybody; ii) offers evidence-based assessments; iii) addresses a variety of user-specific concerns/challenges/problems; iv) draws upon a sound psychosocial theoretical framework; v) draws upon sound counselling approaches; vi) understands and mimics linguistic styles as found in the counselling environment, and finally vii) interprets client sentiments by recognising and interpreting a range of trigger words offered by the client during interaction.

4.3.1.1. Goal 1: Accessible counselling

- Objective 1.1: To develop a framework that can provide counselling to a vast user base;
- Objective 1.2: To develop a framework that can provide counselling regardless of traditional limitations, such as operation hours, scheduling, transport challenges and so on;
- Objective 1.3: To develop a framework that can provide immediate support;
- Objective 1.4: To develop a framework that is convenient and easy to use;
- Objective 1.5: To develop a framework that has little or no financial implications for the user;
- Objective 1.6: To develop a framework that encourages users to access mental health services, which would otherwise have been difficult to access;
- Objective 1.7: To develop a framework that can provide basic mental healthcare services to multiple users simultaneously;
- Objective 1.8: To develop a framework that eliminates the stigma associated with the use of mental health services;
- Objective 1.9: To develop a framework that ensures user comfort;
- Objective 1.10: To develop a framework that is consistent in its services; and

- Objective 1.11: To develop a framework that does not offer its services as the only option but encourages and motivates towards the attendance of in-person sessions, if required.

4.3.1.2. Goal 2: Address user-specific concerns/challenges/problems

- Objective 2.1: To address career and work-related issues;
- Objective 2.2: To address personal development (general issues);
- Objective 2.3: To address dysfunction;
- Objective 2.4: To address emotional wellbeing;
- Objective 2.5: To address family dynamics;
- Objective 2.6: To address interpersonal relationship dynamics;
- Objective 2.7: To address personal development concerns;
- Objective 2.8: To address psychological challenges; and
- Objective 2.9: To address requests for help/support.

4.3.1.3. Goal 3: Evidence based assessments

- Objective 3.1: To develop a framework that provides evidence-based assessments based on the user's specific needs.

4.3.1.4. Goal 4: Draw upon a sound psychological theoretical framework

- Objective 4.1: To utilise solution-focused brief therapy (SFBT) principles.

4.3.1.5. Goal 5: Draw upon sound counselling approaches

- Objective 5.1: To provide advice and support;
- Objective 5.2: To incorporate a sound, commonly used and recognisable psychosocial approach/paradigm as the framework's underpinning approach; and
- Objective 5.3: To incorporate sound therapeutic approaches.

4.3.1.6. Goal 6: Understand and mimic linguistic styles

- Objective 6.1: To understand and mimic informal communication;
- Objective 6.2: To understand and mimic emotional and expressive styles;
- Objective 6.3: To understand and mimic various interaction styles;
- Objective 6.4: To understand and communicate written expressions;
- Objective 6.5: To understand and communicate different tones and politeness;
- Objective 6.6: To understand and use creative language; and
- Objective 6.7: To understand and communicate in formal and specialised language.

4.3.1.7. Goal 7: Recognise and interpret triggers words

- Objective 7.1: To grasp and interpret terminology related to various realms such as emotions, feelings, relationships, family dynamics, self-awareness, personal growth, mental health, wellbeing, work, career life events, challenges, personal habits, lifestyle, social factors, environmental influences, support, and coping strategies; and
- Objective 7.2: According to Deshpande and Warren (2021), the ability of chatbots to recognise emergency circumstances and take appropriate action raises serious concerns. These issues include the possibility of suicide, self-harm and risk to others. Consequently, the framework must be capable of identifying and interpreting trigger words indicative of risks to oneself or others.

The researcher proposes these criteria as critical standards to be adhered to during the development process. This implies that the actual programming must be employed to develop a system that can acceptably perform the above activities.

4.3.2. ANALYSIS PHASE

The second phase (analysis phase) of the SDLC involved working on the root of the problem that had to be solved. During this phase, possible solutions were submitted and analysed to determine the best solution for the project's penultimate goals. During the analysis phase, the functional requirements of the solution were considered by incorporating data from previous analyses. This phase is generally regarded as less theoretical, with a greater emphasis on the practical implications (Khan & Kumari, 2021; Pothukuchi *et al.*, 2023).

To complete the analysis phase, the researcher's focus had to be on the identification of problems/needs, functional requirements, scope, stakeholder involvement, goal alignment, risk assessment, feasibility and documentation. These are discussed in the sub-sections below.

4.3.2.1. Identification of problems/needs

The central aim proposed the development of an AI-driven chatbot counselling framework in social work and EHWP to modernise the field, expand knowledge and maintain relevance in a technologically advanced world. Findings suggested that there is a growing interest and need for an intelligent online counselling framework within the digital ecosystem of an EHWP.

The preceding sections of this study found that mental health chatbots could be a key strategy in embracing digital data and the implementation of digital interventions in the field of social work theory and application. It was found that systems must be developed through the understanding of the market need within the constantly changing state of information needs and demands, taking into account that no single system can meet all organisational needs. However, the adoption and

implementation of technology, which is a long-term strategy, inevitably leads to the improvement of organisational processes (Roestenburg, 2021).

Problems or needs that have been identified are twofold and include the need of the EHWP, as well as that of the client base. Goal 1 addressed the EHWP and user needs, while Goals 2-7 described the user needs.

To propose a workable product, sound research methodologies had to be applied. The applied methodologies, which included a literature study, collection and analysis of live chat transcripts in the form of thematic-, textual- and sentiment analyses, yielded research findings in the form of big data. The processed and analysed data have been used to inform the proposed product.

4.3.2.2. Functional requirements

Functional requirements include both clinical requirements and technical requirements. For this study, only the clinical requirements for the AI counselling framework were discussed. These clinical requirements, termed clinical technical requirements, encompass aspects of both clinical practice and technical application within the scope of this study. The rigorous technical prerequisites exceeded the parameters of this study, particularly within the realms of IT and HIT, which necessitate profound domain-specific proficiency.

4.3.3. AN INTELLIGENT ONLINE COUNSELLING FRAMEWORK: CLINICAL TECHNICAL REQUIREMENTS

This section provides a synthesis of the clinical technical requirements for establishing an intelligent online counselling framework. The section discusses the essential functionalities and considerations required to enable a robust and responsive AI-driven counselling framework. The clinical technical requirements considered includes NLP, emotion recognition, affective computing, contextual awareness and goal-oriented interactions.

In addition, the section discusses a knowledge base, dynamic risk assessment protocols and adaptability in learning. Emphasis is place on therapeutic goal setting, user engagement strategies and evidence-based assessments. The discussion extends to the ethical considerations in AI integration, privacy, data protection and user autonomy. Mobile application integration, administrative alignment and community resource referrals are considered as practical components.

4.3.3.1. Natural Language Processing (NLP)

By utilising one or more advanced state-of-the-art NLP techniques, such as word embedding, sentiment analysis, sequence-to-sequence models and attention mechanisms, the model would ideally have been able to understand and interpret natural language to provide intelligent responses to user input. Ideally, the model would be able to provide appropriate contextualised

responses to complex language input and thereby establish interactive communication with users (Bendig *et al.*, 2022; Bhirud *et al.*, 2019).

The Delphi panel expressed the importance of the user feeling "heard and understood" (Noyce & Simpson, 2016). The panellists added that the framework should be able to exhibit a significant degree of empathy through its responses. An empathic process would include recognising and appropriately responding to subtle emotional cues in user interactions.

Concerns were raised by the Delphi panel regarding the adequate risk assessment abilities of the model. The panellists believed that the framework must incorporate channels through which protective factors, such as values, can be identified and used as motivation. The framework, furthermore, would have to be trained to identify high-risk situations and would need to refer these users to human counsellors when appropriate and necessary.

Delphi panellists further noted the importance of session continuity, whereby the framework can remember and recall past interactions and can continue where the user left off. The persistence context of the framework is essential for building a sense of ongoing support and continuity in the therapeutic process.

Goldberg *et al.* (2020) provide six practical suggestions for the development of counselling models in their study on NLP and ML. These are:

- ML is most promising when observing and predicting observable linguistic behaviours;
- ML models are best trained by human coding;
- ML models are to be trained using large data sets;
- Develop ML models using training data sets and test models using test sets;
- Interdisciplinary collaboration needs to take place between clinicians, researchers with expertise in processing linguistic data and professionals with backgrounds in computer sciences; and
- Expectations should be reasonable, especially concerning when these models will become viable tools.

Specific state-of-the-art NLP techniques, such as transformer-based models (for example Bidirectional Encoder Representations from Transformers [BERT] and GPT) excel in contextual comprehension, can handle complex language inputs and can understand and generate human-like responses (Acheampong *et al.*, 2021; Yildirim & Asgari-Chenaghlu, 2021). However, a significant obstacle highlighted by the Delphi panel surrounds the framework's ability to detect when a user is not being truthful. In traditional settings, therapists rely on observations and clinical interpretations to discern dishonesty, whereas the proposed framework would lack such nuanced capabilities. This limitation emphasises a need to integrate capabilities that could flag potential dishonesty based on linguistic cues and patterns.

Lyra SA has a vast number of live chat transcripts available that could be used for the observation and prediction of linguistic behaviours. Empirical research on a sample (elaborately described in Chapter 3) found the following linguistic trends, as displayed in Table 14, which could be of use in the training corpus of the framework's NLP modelling.

Table 14. Linguistic styles

Linguistic styles	Themes	
Informal Communication	Casual communication	Colloquialism
	Emoji use	Informal
	Internet slang	Slang
	Abbreviation	Laughing
Emotional and Expressive	Descriptive language	Positive language
	Emotive language	Reflective language
	Emphatic	Sarcasm
Interaction Styles	Clarifying questions	Open-ended questioning
	Direct questioning	Rhetorical questions
	Inquisitive style	
Written Expression	Errors (typos/misspellings)	Exclamation marks
	Use of punctuation	Paraphrasing
Tone and Politeness	Friendly tone	Polite tone
	Impersonal	Relaxed tone
	Neutral tone	Respectful tone
Creative Language	Analogy	Emoji
	Common expression	Linguistic trends
Formal and Specialised Language	Contraction Usage	Specialised knowledge/terminology
	Formal language	
Miscellaneous	Informative language	Vulgar language
	Personal language	Linguistic trends
	Personal pronouns	Specialised knowledge/terminology

Excerpt 18 [1(7)] below illustrates a user using abbreviations, slang and feelings or emotions in quick succession. The user discussed relationship issues with their previous partner, which they referred to as "my x"; the user then referred to another person, presumably their current partner, as "M", whereafter they referred to another person (presumably their previous partner) as "R". The user proceeded to use a slang terminology, "tipping", which possibly refers to the dislike or aversion of something. In the context of this session, it seems as if the user had an aversion to their current relationship status. The user immediately moved to feelings or emotions, noting that they feel depressed, ascribing the emotional experience to their job and salary.

The significance of the excerpt illustrates how complex it can become for the model to understand and interpret the message. In addition, the model would have to understand the message's intent, whilst calculating an applicable response.

User	*So R who my x wanted me back, because M is not a guy I see a future with.*
User	*But then I realize that I don't want be in a relationship with R and leave M like that cause he was there for me.*
User	*So my point is I have a problem*
User	*I don't know how to handle the situation anymore*
User	*There is a lot more then what I said*
User	*Just hate tipping*
User	*I feel depressed*
User	*my work makes me feel depressed*
User	*my salary makes me feel depressed*

Excerpt 18. Example of user linguistic style

4.3.3.2. Emotion recognition and affective computing

Emotion recognition and affective computing refers to a system's ability to recognise human emotion, sentiment and feelings and the ability to accurately respond to these (Wang *et al.*, 2022). Affective computing has been successfully applied in various fields, such as education, marketing, gaming, robotics, autonomous vehicles, virtual reality, security, surveillance and customer service to respond to human affect. However, emotion recognition and affective computing are crucial components of any counselling model. For dialogue systems to have a therapeutic conversation with a user, they must be able to correctly recognise a user's emotional state. Oh *et al.* (2017) have found that accurate emotion recognition provides better user satisfaction, lending itself to better therapeutic outcomes.

Empirical research conducted in Chapter 3 on emotion or sentiment involving live chat counselling transcripts has yielded several results. These are displayed in Table 15.

Table 15. Empirical findings of sentiment analysis

Sentiment	Count of sentiment
Sentiment: Negative	3030
Sentiment: Neutral	1639
Sentiment: Positive	1809
Grand Total	6478

Numerous emotion recognition algorithms exist. These algorithms differ in input channels, output labels, representation models and classification methods (Landowska, 2018). Landowska (2018) notes that there are three types of emotional state representations, namely discrete-, dimensional-, and componential representations.

- A discreet model distinguishes a set of basic emotions or word labels and label representations as belonging to an emotional state based on predefined criteria;

- A dimensional model refers to a point in a multi-dimensional state. The circumplex model of affect is informed by the dimensional model, whereby any emotion might represent a point in a space of two continuous dimensions of valence and arousal - this dimension differentiates positive from negative emotions. This type of model takes six emotions into account;
- A componential model draws upon several factors that constitute or influence the resulting emotional state. The model defines a hierarchy of 22 emotion types, representing various emotional states that might be experienced. In contrast to the other two models, the componential model takes the generation of emotions into account.

The use of a componential model suited the model under investigation well. It was used to identify various mood states and to respond accordingly to the user (Figure 33).

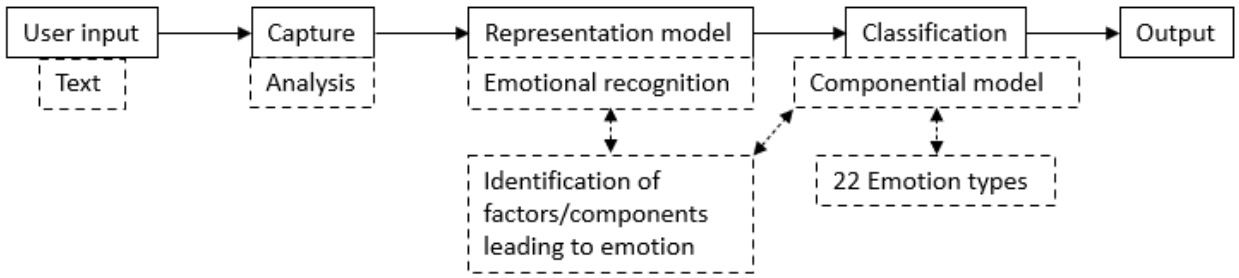


Figure 33. Componential model

The data indicated in Table 16, as derived from the empirical data (emotive expressions occurring >0.01%), have been identified as emotional expressions and can serve as an example of the necessary approach to ensure accurate affective computing (please note that the data were not extensive and only serve as an example). The emotions are categorised within the table according to sentiment analysis previously done and not according to any theory of emotion.

Table 16. Emotive expressions occurring >0.01%

Sentiment	Word	
Negative	Anger	Divorce
	Angry	Cheating
	Negative	Anxious
	Overwhelm	Overwhelmed
	Stressor	Panic
	Stressors	Fear
	Pressure	Anxiety
	Challenge	Stress
	Hate	Stressed
	Hurt	Struggle
	Adversity	Sad
	Struggles	Alone
	Abuse	Sadness
	Abusing	Cry
	Abusive	Depressed

	Problem Bad Difficult Struggling	Depression Guilt Guilty
Positive	Positive Hope Adapt Recovery Accomplish Accomplished Achieve Well Support	Care Great Friends Healing Grow Proud Happy Glad Happiness

As evident from the data, which was analysed using AI sentiment analysis, it is not a simple task to categorise emotive words. This is because some emotive expressions might fit into multiple categories. It is, therefore, important that the framework considers the context of the discussion in its entirety to make an accurate sentiment analysis.

Excerpt 19 [1(58)] illustrates a user most likely experiencing work-related and personal stressors, without pertinently referring to stress. The user noted stressors due to restructuring at their workplace. The user continued, stating that they experienced physical symptoms, which they ascribed to tension (presumably stress/anxiety symptoms). The user elaborated that they had limited time for physical activity, whereafter they referred to stressors in their personal life.

The significance of this interaction for affective computing or emotion recognition is seated in the fact that the user did not pertinently refer to stress/stressors/anxiety. Thus, the framework would have to search for the significance of emotion in the message, without it being stated.

User	*So managing product development, but also have an Operations team and a Training/product support team in my structure, dealing with the day-to-day running of the business*
User	*Recent restructuring in the bank is also adding to the additional pressure*
User	*My struggle with tension headaches, sore back and shoulders...*
User	*Hardley any time for physical activity*
User	*Also a mom... 2 adult kids, a 14yr old and an 11 year old. Last two are my step children, but I take them as my own. Married their dad who was a widower nearly 7 years ago*
User	*Soon to be grandma... any time over the next two weeks*
User	*I am 51 years old*

Excerpt 19. User emotive expressions

4.3.3.3. Contextual awareness

Contextual awareness refers to the framework's awareness of the context, namely, linguistic-, persistence- and goal context, within which it would function. This could be achieved by implementing techniques such as dialogue state tracking, which allows the framework to remember, recall and refer back to previous sections of the conversation, thus ensuring

coherence and relevance (Wang, 2023). The framework would have to understand the words and phrases that are used when interacting with the user. This could be achieved by the ML through training sets used to train the framework in different contexts.

Persistence context refers to the framework's ability to remember and retrieve information within the conversation that it is currently having with the user. The framework would, therefore, not chat sentence-to-sentence but converse within the context of the entire conversation (Koulouri *et al.*, 2022; Shinozaki *et al.*, 2015). The framework would have to consider the entire conversation when conversing with a user because:

- Users use multiple sentences across a space of time to describe the challenge/issue;
- Users move back and forth between challenge descriptions and challenge impacts;
- Users move freely between challenge descriptions, ideal futures and goals;
- Users use relating and unrelating sentences in the same conversation; and
- Users connect with multiple challenges.

Understanding the challenge and the related goals within the context of the conversation would be essential for the framework's ability to remain goal-orientated within the entirety of the conversation. This, in turn, would allow the framework to align with the user's goals and aspirations, as illustrated in Figure 34.

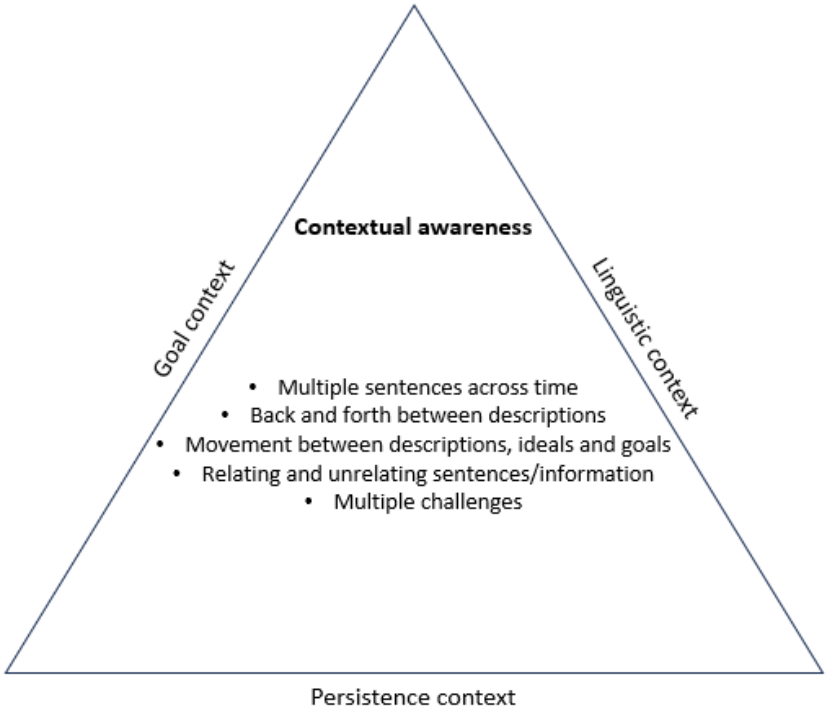


Figure 34. Contextual awareness

Excerpt 20 [1(70)] illustrates a user referring to multiple challenges spanning past and present experiences. The user made note of past trauma (without specifying the trauma), whereafter they immediately moved to the present, referring to present experiences, specifying anxiety. The user then moved on to discuss the impact of the mentioned challenges on their occupational performance. The user continued by discussing current beliefs relating to their confidence. Throughout, the user did not respond to the agent's inputs or questions, necessitating the agent to engage with the user within the context that the user created.

The excerpt illustrates how a framework should be able to consider the goal context, the linguistic context, the descriptive context and the intention context between various periods within a persistence context. The persistence context would, therefore, be essential for the framework to comprehend the message accurately and to compute a relevant response that proves the framework's contextual awareness.

User I think it's a combination of many past traumas. Yesterday, I was doing a work assignment and I volunteered to be the one who will present our discussion and that really didn't go well. I literally had my notes and had to just present what I wrote down. But I got nervous, I stuttered and I missed some of the things we discussed because my anxiety just took over. And i have been trying to push myself to talk more in meetings and really practice speaking out more and this one time I volunteer to do so I messed it up and I have been beating myself up for it

Agent I can see how hard you have been trying. I have to comend you for that. I think it is very natural and a valid response to feel nervous and anxious in these social situations. .

Agent What motivated you to put yourself out there? To confront this fear? .

User I also feel like it's affecting my work though as I work in a fully virtual company and have these meetings weekly. It's easy to get lost and not be memorable and because of my role in the company I feel that I should find a way to at least feel comfortable talking to people. It's one of the things I feel that it's lacking in my job role, just being able to articulate myself properly in meetings

Agent I think that fact that you are aware of this and attempting to enhance this skill is fantastic. It requires willingness and courage. But let's explore a bit more. When you find yourself feeling anxious in these situations, where are your thoughts at? What are you thinking prior and during this time that might add to the anxiety?.

User I always think I'm gonna mess up. That I'm not good enough, smart enough. I've really had a hard time believing that I deserve this job in the beginning because of the job spec vs my skills and then I killed the interview and I hit the job and even though my manager has assured me that I'm doing amazing and I really was the best out of the bunch of applicants it was hard for me to believe it you know.

User I also think because I work with people who have big personalities and are really smart and hardworking. maybe I try to compete with that

Agent I completely understand what you are saying. But is seems to me that you thoughts and fears of not being good enough is prohibiting you from acting in accordance with your capabilities. Your thoughts regarding yourself is at this time your biggest stumbling block. Comparing yourself to others is also not helpful. .

Excerpt 20. Contextual awareness

The implementation of a robust DM system that can manage complex interactions and context effectively, such as Rasa or Microsoft's Bot Framework, should ideally be implemented (Singh *et al.*, 2019). The framework should maintain and retrieve contextual information across various sessions from within a securely stored history, which would enable the framework to provide personalised experiences.

4.3.3.4. Knowledge base integration

The framework's knowledge base can be described as the framework's database, which serves as a comprehensive repository of information. It contains information on challenges, counselling approaches, therapeutic techniques and resources, providing the framework with the necessary knowledge to understand, assess and address the needs of service users. The framework should draw upon this knowledge base and interpret input, correlate it with available data and develop an output. Baez *et al.* (2020) recommend schema mapping, which integrates data from various sources.

Dynamic knowledge bases that can be updated in real time should ideally be developed. Graph databases such as NEO4j would be ideal for managing complex relationships between varying sources of information. Schema mapping techniques can be utilised to seamlessly integrate data from various sources. This would ensure that the framework could access a wide range of information to assist with user queries (Agrawal *et al.*, 2023; Kejriwal, 2022).

The knowledge gained from empirical research must be integrated into the framework's knowledge base. In the context of this study, extensive insights have been developed into users' reasons for engaging in live chat counselling sessions with human counsellors. The empirical study's data indicate that users engage for several reasons, as shown in Table 17 below.

Table 17. Reasons for user engagement

User need(s)	Frequency
Personal development: General	305
Dysfunction	274
Psychological challenges: Uncategorised	228
Emotional wellbeing	196
Career and work-related: general	174
Interpersonal Relationships	104
Family dynamics	101
Personal development: Encouragement	59
Personal development: Empathy	39
Dysfunction: Uncertainty	22
Stress	19
Separation/divorce	17
Personal development: Gratitude	14
Personal development: Self-Care	13
Personal development: Encouragement, Personal development: Positive feedback	12
Fatigue: Uncategorised	11
Feeling overwhelmed	11
Interpersonal conflict: partner/spouse	11

Not coping	11
Psychological disorder: Suspected	11
Dysfunction: Uncertainty, Dysfunction: Questioning	10
Request for help/support: Suggestion	10
Experience of unpleasant emotions	9
Loss/bereavement	9
Anger	8
Emotional fatigue	8
Personal development: Self-improvement	8
Request for help/support: Advice	8
Childcare	7
Personal development: Reassurance	7
Interpersonal communication	7
Suicidal ideation	7
Religion	6
Trauma: Various	6
Dysfunction: Uncertainty, Dysfunction	6
Emotional wellbeing: Anxiety	6
Family dynamics: Family	6
Personal development: Positive reinforcement, Personal development: Encouragement	6
Abuse	5
Compassion fatigue	5
Dysfunction: Frustration	5
Dysfunction: Questioning	5
Emotional regulation	5
Emotional wellbeing: Anxiety, Dysfunction: Uncertainty	5
Family dynamics: Family relationships	5
Infidelity	5
No support system	5
Parental guidance	5
Career and Work-related Issues: Time management	4
Domestic violence	4
Emotional wellbeing: Mental health	4
Feeling “down”	4
Financial Issues: Financial stress, Family dynamics: Family issues	4
Personal development: Empathy, Request for help/support: Supportiveness	4
Personal development: Encouragement, Personal development: Motivation	4
Personal development: Encouragement, Personal development: Self-care	4
Personal development: Encouragement, Request for help/support: Suggestion	4
Personal development: Motivation	4
Personal development: Resilience	4
Physical abuse	4

Verbal abuse	4
Request for help/support: Problem-Solving	4
Emotional abuse	3

The listed reasons represent just a fraction of the potential reasons why a user would use a chatbot for counselling. However, the table does provide an accurate representation of frequent needs for engagement.

Excerpt 21 [1(85)] illustrates a user referring to themselves as a “control freak”. The framework would need to be able to cross-reference this term in its knowledge base to determine the relevant context to which it refers. The user continued to indicate that they are receiving treatment to conceive a child but also noted that the treatment might not be effective. The framework should understand (within context) that the user is referring to challenges relating to infertility and the accompanying emotional responses. The user continued to note that they prioritise work instead of their wellbeing, which would necessitate the framework to have knowledge of work-life balance. The user continued to shift their focus to the workplace, stating that they would need support from their place of employment, without needing to divulge too much information to their employer. The user finally asked for guidance on how to remain focused on their job while experiencing personal stressors. The excerpt reiterated the need for the framework to have a knowledge base of various situations and contexts, whilst referring to the knowledge base to compute an appropriate response.

User "It is a highly challenging situation, given that I am a bit of a control freak, and had only planned on starting a family in the next 2 to 3 years.

I have been advised by my doctor that I would need to proceed with treatment and conceive a child within the next year.

It is likely that even with treatment, that I am still unable to conceive.

I am struggling with the shift in focus from my job being my priority to my health and wellbeing needing to be the priority.

This illness has become all consuming in a matter of weeks, and it is difficult to reconcile the idea of what I had planned for my life vs the reality and urgency of the situation. "

User "I am also going to need significant support from my manager and colleagues once this treatment process kicks off in earnest in January 2023.

I have had discussions with my line manager to explain the diagnosis and impact without divulging too much information, as I believe there needs to be a separation of work and personal life.

I am feeling uncomfortable with explaining too much of my medical space to my manager, as I do not want to be disadvantaged regarding opportunities in the work place because of this illness that is out of my control.

I do not want to be viewed as a non performing member of the team, but I am struggling to be as deeply invested in my job as I was before the diagnosis.

Please can you guide me on how to remain focused on work while my personal life is in disarray. "

Excerpt 21. Need for extensive knowledge base

4.3.3.5. Dynamic risk assessment and crisis response

The framework should be able to identify risks and crises, such as suicidal ideation, homicidal ideation, self-harm, substance use/misuse and all forms of abuse. A thorough risk assessment would need to be completed using validated risk assessment tools and protocols. All risks should be reported to a human counsellor for follow-up and follow-through interventions (Martinengo *et al.,* 2022).

According to the Delphi panellists, the framework must be adept at identifying and reporting any forms of abuse, given the ethical implications. The framework should ideally integrate thorough risk assessment protocols, whereafter positively identified risk situations should be reported to human counsellors.

As per the Delphi panellists, ethical considerations are paramount in all risk scenarios. The framework should, therefore, be able to navigate the balance between user privacy and ethical responsibility in an attempt to prevent harm. The reporting of abuse involves various ethical implications, which are aimed at ensuring user confidentiality while adhering to legal and ethical mandates for reporting.

Further risks and ethical considerations may be associated with the assumption that a framework could interpret colloquialisms or sarcasm in the wrong way and produce false positives or negatives in crises. The urgency, therefore, on an ethical basis to do no harm necessitates strong mechanisms to detect and correctly handle these nuances, making sure that no possible harm is extended to the user, either emotionally or psychologically.

In the context of dynamic risk assessment and crisis response, the Delphi panellists are of the opinion that the framework must consider various legislative frameworks regarding abuse. Various acts and legislation are aimed at protecting vulnerable individuals. The framework must be aware of such legal parameters.

The empirical research has highlighted the risk codes displayed in Table 18. These risks are by no means extensive but do provide an indication of which risks users often report.

Table 18. Risk codes

Risk	Count
Emotional wellbeing: Suicidal ideation	2
Emotional wellbeing: Suicidal ideation, Emotional wellbeing: Depression	2
Dysfunction: Abuse, Dysfunction: Defensiveness, Dysfunction: Blaming others	1
Dysfunction: Addiction	1
Dysfunction: Addiction, Dysfunction: Unreliability	1
Dysfunction: Addiction, Personal development: Personality Change	1
Dysfunction: Addiction, Personal development: Self-awareness	1

Dysfunction: Addiction, Personal development: Self-awareness, Psychological challenges: Resistance to change	1
Dysfunction: Addiction, Personal development: Self-improvement	1
Dysfunction: Alcohol abuse, Interpersonal Relationships: Relationship difficulties	1
Dysfunction: Alcohol consumption	1
Dysfunction: Alcohol consumption, Family dynamics: Family relationships, Interpersonal Relationships: Romance/dating	1
Dysfunction: Alcohol use	1
Dysfunction: Alcohol use, Emotional wellbeing: Coping mechanism	1
Dysfunction: Alcohol use, Interpersonal Relationships: Relationships	1
Dysfunction: Alcoholism, Family dynamics: Family stress	1
Dysfunction: Domestic abuse, Request for help/support: Requesting assistance	1
Dysfunction: Domestic violence, Family dynamics: Family conflicts, Dysfunction: Double standards, Emotional wellbeing: Emotional abuse	1
Dysfunction: Domestic violence, Family dynamics: Family issues, Family dynamics: Childcare	1
Dysfunction: Drug abuse	1
Dysfunction: Drug use	1
Dysfunction: Drug use, Dysfunction: Harm	1
Dysfunction: Drug use, psychological challenges: Lack of communication, Dysfunction: Frustration, Interpersonal Relationships: Relationship issues	1
Dysfunction: Drugs, Career and Work-related Issues: Legal issues, Financial Issues: Financial problems, Career and Work-related Issues: Health issues	1
Dysfunction: Physical abuse, Dysfunction: Verbal abuse, Dysfunction: Blame, Family dynamics: Parenting	1
Dysfunction: Physical abuse, Interpersonal Relationships: Supportive relationships, Family dynamics: Family relationships, Emotional wellbeing: Stress	1
Dysfunction: Physical violence, Financial Issues: Financial constraints, Career and Work-related Issues: Transportation difficulties	1
Emotional wellbeing: Alcohol addiction, psychological challenges: Lack of social support, Family dynamics: Family conflict	1
Emotional wellbeing: Emotional abuse, Dysfunction: Power imbalance, Family dynamics: Family conflict	1
Emotional wellbeing: Emotional abuse, Interpersonal Relationships: Breakup recovery, Dysfunction: Resentment	1
Emotional wellbeing: Emotional abuse, psychological challenges: Fear of judgment, Interpersonal Relationships: Personal relationships, Personal development: LGBTQ+ identity	1
Emotional wellbeing: Substance abuse	1
Emotional wellbeing: Substance abuse, Dysfunction: Domestic violence	1
Emotional wellbeing: Substance abuse, Dysfunction: Self-blame	1
Emotional wellbeing: Substance abuse, Family dynamics: Family dysfunction, Family dynamics: Child neglect	1
Emotional wellbeing: Substance use disorder, Emotional wellbeing: Negative coping mechanism, Emotional wellbeing: Mental health, Dysfunction: Addiction	1
Emotional wellbeing: Suicidal ideation, Emotional wellbeing: Anxiety	1
Emotional wellbeing: Suicidal ideation, Emotional wellbeing: Depression, Dysfunction: Fatigue	1
Emotional wellbeing: Suicidal ideation, Emotional wellbeing: Mental distress	1
Emotional wellbeing: Suicidal ideation, Emotional wellbeing: Mental health concern	1
Emotional wellbeing: Suicidal thoughts, Emotional wellbeing: Depression	1
Emotional wellbeing: Suicide ideation, Emotional wellbeing: Past suicide attempt, Emotional wellbeing: Current suicidal behaviour	1
Emotional wellbeing: Suicide prevention, Personal development: Empathy	1

Trigger words, as they appeared in the empirical research, indicate that risk protocols should be activated once the framework detects them. These include the words listed in Table 19 (the list is based on the sample and is by no means extensive).

Table 19. Risk trigger words

Cluster and cluster frequency	Trigger word and word frequency	
Emotions and Feelings (Total: 5.96%)	sad (0.02%) anxious (0.03%) overwhelmed (0.03%) anger (0.03%) guilt (0.01%) numb (0.01%)	suicidal (0.01%) suicide (0.01%) panic (0.01%) depressed (0.02%) hate (0.01%) hurt (0.04%)
Mental Health and Wellbeing (Total: 1.60%)	depression (0.02%) bipolar (0.01%) anxiety (0.06%) trauma (0.01%)	addiction (0.01%) alcohol (0.01%) smoking (0.01%) medication (0.01%)
Personal Habits and Lifestyle (Total: 0.10%)	addiction (0.01%) smoking (0.01%) drinking (0.01%)	sleep (0.01%) self-medicating (0.01%)

As soon as the risk protocol is activated, a referral to a human counsellor should be made to ensure immediate and comprehensive intervention. The envisaged referral process and human intervention are illustrated in Figure 35.

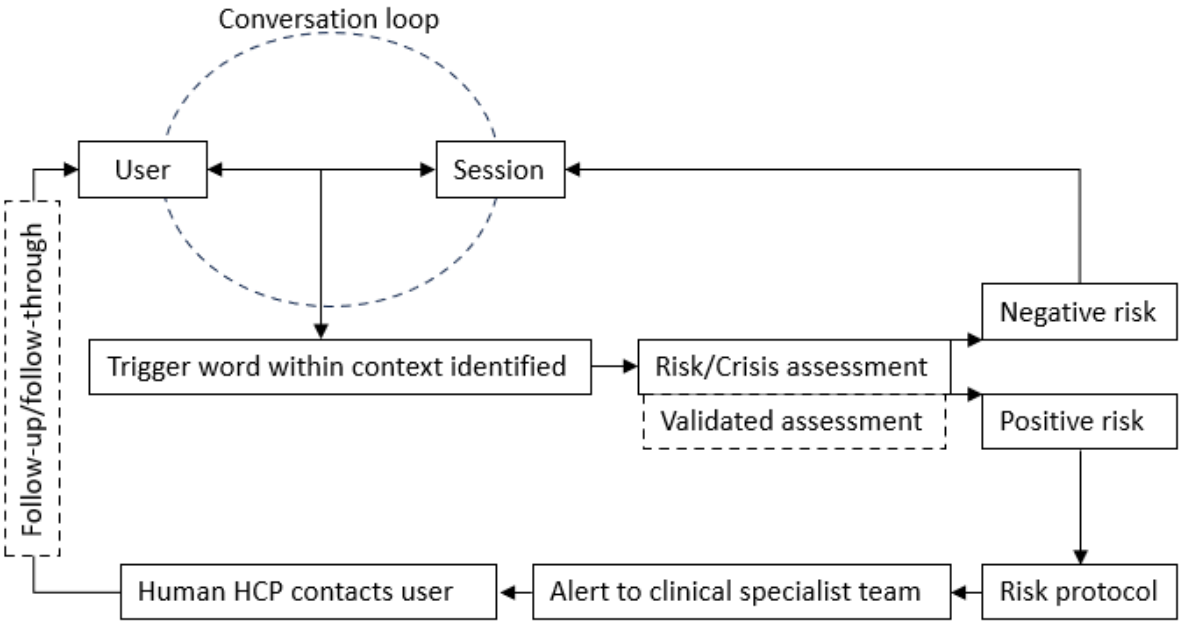


Figure 35. Dynamic risk assessment and crisis response

Excerpt 22 illustrates [1(47)] a potentially suicidal user connecting for assistance, who stated that they might be suicidal and need to talk to someone. The framework should be able to identify words and phrases such as “suicidal”, “feels too much”, “life would be better without me”, “if I die”, and so forth as risk utterances and should trigger an alert with the relevant department for a

human counsellor to make contact with the user. It would, therefore, be beneficial if the framework could conduct a risk assessment while the client is engaged.

User	*i just really need to talk to someone*
Agent	*Hi XXX.*
Agent	*Thank you XXX , what is overwhelming you if i may ask?.*
User	*i think i might be suicidal*
User	*things have taken a lot of turns in my life and recently it just feels too much*
User	*and im having vivid thoughts about how life would be better without me*
User	*and that people actually might care for 2mins if i die*
User	*i know it sounds dramatic and stupid. but thats how i feel*
Agent	*Would you like me to request a counsellor to give you a call? .*

Excerpt 22. Risk engagement

4.3.3.6. Adaptability and learning

Ahmad *et al.* (2022) propose an established personality model with advances in NLP. The authors suggest the notion of an adaptive personality model, which can recognise personality traits of users and in turn, enable them to adapt to the changing needs and states of users when establishing personalised interactions. The framework would ideally be able to adopt its responses based on the user's input, within the confinements of the user's emotional and contextual experience. The framework should also be able to adapt to circumstances as they change during the sessions and accurately respond to new and additional user information.

Adaptability and learning are essential for the framework, as users (U) and agents (A) do not follow a linear pattern when engaging in live chat counselling. Empirical evidence suggests that sessions can follow any pattern of interaction, such as U; A; U; A or U; U; U; U; A; A (amongst others). Problems or challenges are discussed in the same way, namely P1; P1; P1 or P1; P2;P1 or any other pattern. It is essential for the framework to adapt to the user's conversation style and to remain within the context of the user's needs.

Excerpt 23 [1(74)] illustrates a user's request that previous session details be recalled, to remain within the context of previous sessions. The user changed their emotional experience and context frequently during the engagement, with the agent adapting their approach to each of the changes. At this stage of the engagement, the framework must be able to recall previous session information within the context of previous sessions and incorporate previous session information into the current session, while considering the personality traits of the user. The user in Excerpt 23 changed the context of the discussion and their emotional experience 19 times in 13 engagements. The user expressed feelings of guilt, minimised their own experience, considered ending the session, explained intentions, mentioned hurting, conveyed feelings of loneliness,

expressed concerns regarding counsellor care, and even minimised themselves to only being a patient. They also focused on the agent as a person, expressed moral convictions due to using counsellors, expressed concern for other persons using the service, expressed discontent with engaging with a different counsellor during each session, and raised concerns regarding a lack of a therapeutic relationship with one counsellor. Furthermore, they felt that the therapeutic relationship was one-sided, felt as though they were using the counsellor, viewed the counselling relationship as a toxic relationship, expressed dissatisfaction with the previous session, expressed gratefulness and sought advice.

User Can you read the messages from earlier? Can you help me with that?

Agent Hi XXX, please give me a moment to read through your previous chat.

User oki doki

Agent Okay, I see that you mention that you have been feeling guilty when not talking to us but also feel that it would be a waste of our time of you chatted with us without there being a problem....

User yes, but this was a problem because ive been thinking about ut for a good while and its making me feel bad. if you dont see it as a big enough problem, thats okay. im happy to leave.

User i dont mean to sound mean... sorry

Agent Oh, I see what you mean. Why do you feel it is a problem for you to feel that way? .

User Because I am hurting alone. You guys dont ever think about me. The last time you guys think about me, is when my last message is sent. Not a minute after that. I hurt while the people I hurt for never even think about me

User Why should I hurt while you guys never even feel anything. Im just a patient to you guys. But youre not just therapists. Youre people. And morally I feel wrong when just using you to feel better

Agent What gives you that impression XXX? What makes you believe that we forget about you as soon as you leave this platform?.

Agent Do you really feel like you are using us? Remember this platform was created for you to reach out whenever you need support. Yes, we are people, but we are also here to provide you with the support that you require.

User Because im not the only patient. Im not the only person you help. Im not the only person hurting. And every time I talk to someone, its someone new. We only interact for more or less 1 hour and then were done. Why would someone think about me or worry about me when weve never bonded. Never actually talked. They never even tell me their names, so why would mine matter? You guys get a different name every time. I never get a name so they become one person. Therefore, I feel more connected to you than you do to me.

User Yes i feel like Im using you and I do know its your job. But youre doing the thing that most of you spend the most time on: Getting people out of toxic relationships. And this is one. I only ever come to you to use you, and you just oblige. That doesnt sit right with me

Agent Makes sense. I hear what you are saying and get where you are coming from. You end up feeling like you are connecting more with us than we are with you?.

Agent I can tell that it doesn't sit well with you .

User Yes... and telling me to just not feel bad, wont make me stop feeling bad and just cure me

User I know I came across a bit harsh to the previous guy, but he was essentially telling me, while im drowning, to just stop. So, i got frustrated

Agent Oh yes, no I can see where your frustrations are coming from. You do not want to be told not to feel bad whereas you are feeling bad.

User Exactly

User But you seem more human, so thank you

User So what do I do? How do I stop?

Excerpt 23. Adaptability and learning

According to Garcia (2024), personalised adaptive learning algorithms that can tailor responses based on individual user interactions and preferences should be implemented. Techniques such as reinforced learning would optimise the framework's response over time. Feedback loops where users rate the quality of feedback would contribute towards the accuracy of the framework's performance. ML pipelines - which enable continuous improvement and update the framework based on new data (Steidl *et al.*, 2023) - could ensure that the framework evolves and stays relevant over time.

4.3.3.7. Therapeutic approach, goal setting and progress tracking

The SFBT paradigm is Lyra SA's preferred theoretical framework and therapeutic approach. Thus, all the data used in the analysis and creation of the suggested framework were based on SFBT concepts. It should be evident that the suggested product would operate within an SFBT paradigm because all training and data sets were associated with SFBT.

There is theoretical validity for the suggested framework to adopt the SFBT approach, as it could enable the framework to identify different approaches to understanding a user's concerns based on the semantic combination of words used in the challenge descriptions. By analysing the meaning embedded in user language, the framework could focus on identifying solutions and leveraging the strengths of the user, rather than focusing on the challenge itself.

The empirical findings suggested that users would often engage in chat services to consult regarding problems that have been grouped in cluster classification in Table 7. It can be assumed that users would make use of the framework for the same or similar reasons. It would also be reasonable to suggest that the framework be trained with techniques in addressing the problem clusters as indicated in Table 20.

The Delphi panellists believed that the framework could be effectively employed for homework, psychoeducation and skills development. Instead of relying on social media platforms, such as WhatsApp, users could benefit from the guidance and support of the framework. Users could refer to the framework for guidance and support, thereby enhancing their learning and coping strategies through structured interactions and facilitated goal setting.

Table 20. Problem clusters

Problem cluster	Frequency of cluster
Career and Work-related Issues	174
Dysfunction	274
Emotional wellbeing	196
Family dynamics	101
Financial Issues	43
Interpersonal Relationships	96

Personal development	305
Psychological challenges	228
Request for help/support	193
Grand Total	**1610**

The strengths-based approach of SFBT works well for a variety of treatment settings, presenting issues, and patient demographics within crisis care. As a catalyst for change in the SFBT counselling process, a mental healthcare provider and client work together to critically assess the client's subjective perceptions of reality and create new or modified versions of them. Since SFBT is a time-limited counselling modality, the main emphasis is on finding constructive solutions to challenging behavioural issues rather than on the problems' root causes, shortcomings or restrictions.

The proposed product corpus should integrate the most prominent SFBT principles. These include (De Jong, 2015; Finlayson *et al.*, 2021; Johnson *et al.*, 2015 Wulf, 2020):

- Clients are competent in co-constructing goals;
- Clients are the authorities on their own lives and the significance of their experiences;
- There may not always be a link between the problem and the solution;
- Clients must take different actions for change to occur; and
- Change starts with small changes, after which larger changes could follow.

The formulation of goal-oriented questions is one of the additional crucial factors in SFBT. Following a brief description of the issue or circumstance by the client, the co-construction of goals begins with the following questions:

- What would be different once the issue is resolved?; and
- By the end of the session, what needs to change for you to say that our time together was beneficial?

As a foundation for a sequence of follow-up questions and answers, the miracle question (a question guiding clients to envision life if their core issue were resolved) - possibly one of the most distinctive aspects of SFBT - will be essential to the suggested product's probing and solution-generating capabilities. In addition, the following queries should be taken into account in the proposed product's answer:

- Exception-finding questions;
- Scaling questions;
- Coping questions; and
- “What’s better” questions.

Therapeutic approaches employed by live chat counsellors, as identified in the empirical phase of the study, are displayed in Table 21. These approaches should ideally form part of the framework's training set in order to approach the user in a similar manner as a live chat counsellor.

Table 21. Therapeutic approaches

Therapeutic approach (cluster)	Technique	
Cognitive Restructuring and Skill Development	Challenging negative thought patterns Challenging thoughts Distorted thinking patterns	Cognitive reframing Constructive critique Reframing
Crisis Management and Intervention	Crisis intervention Risk assessment	
Client-Centred Approach and Empowerment	Client is owner of challenge Facilitating change Empowerment Identifying goals Identifying strengths Exploration of options	Explore coping strategies Focus on strengths Self-affirmation Self-care What is needed for improvement
Communication and Interpersonal Skills	Assertiveness Empathy Encouragement Validation	Improve communication Open-ended question Paraphrasing
Psychoeducation and Knowledge Sharing	Psychoeducation Interpretation Explanation of thought processes	Homework Normalising Session summary
Support and Rapport Building	Affirmation Offers support	Gratitude Honesty
Problem-Solving and Coping Strategies	Exploration Exploration of thought processes Exploration of options	Explore coping strategies Proposed coping mechanisms Boundary setting Control the controllables

Ideally, the framework would integrate SFBT principles, crisis care and therapeutic approaches with their distinct techniques into its therapeutic approach, goal setting and progress tracking. This would facilitate discussions while guiding the user to tangible solutions to their concerns.

Excerpt 24 [1(27)] illustrates the effective use of SFBT techniques by a live chat agent. The agent highlighted the importance of goals and hope and emphasised clear, achievable goals. Negative thoughts were identified and challenged by drawing on the technique of reframing, highlighted by examples of alternative thoughts. Small steps towards behavioural change were discussed, while emphasis was placed on the ideal future. Scaling questions, while not explicitly used, assisted the user in identifying where they were on the spectrum of their feelings and thoughts. The agent

assisted the client in externalising the challenge to gain perspective. Self-compassion and care were also emphasised by the agent. The user's emotions were acknowledged and validated by the agent, while the agent emphasised the importance of emotional expression in the therapeutic process. The agent finally reinforced the availability of support, which emphasised a collaborative and supportive nature. The framework should, therefore, ideally draw on similar (and other) techniques to provide effective counselling sessions based on the principles of SFBT.

Agent "It is wonderful! We have a goal and a hope! And it is possible. We know have to start taking small steps in pursuit of that.. Being positive means that we have to identify what we are thinking and the way we are thinking and actively challenge those thoughts with an alternative thought. . Example:. ""I am failing in my work."". Alternative:. "" I am doing my best now and learning a lot when I make mistakes"" . "

Agent Does this example make sense? Thinking differently can change the way that we feel and respond. Can you see how this might be very helpful?.

User Yes exactly - how do I get to the alternative

User I can see it and I want to be like that

User I want to be scared to die and not hope that everything kills me and then I didn't technically really give up

Agent 1. Become aware of what you are thinking. . 2. Flip it around. Try to see the opposite to that thought. . . Let's challenge that thought: .

Agent I want to be scared to miss out on life, I never want to give up on me living and experiencing everything that my life can offer. I want to life a full life.

User That is exactly what I want

Agent The do it for yourself. I want to challenge you to write down those thoughts and replace them with alternatives. If we change our perception our feelings and responses change. .

Agent Can I ask you a strange question?.

User Yes you may

Agent Let's pretend that you have grown to be a very old lady. Very old! .

Agent You have lived a full live, there was challenges along the way, but you lived it to the fullest. .

User Yes…

Agent Reflecting back on your life as an old lady, what will make you feel content in the way you lived your life?.

User Hmmm… I finally left my job and went to health sciences and was a nurse or a therapist and I've taken care of people who are at their most vulnerable so it didn't feel like a job. I'm married to my boyfriend and we have successful children. I was able to spoil my siblings the way I've always wanted to. I have lived my life without worrying about money - not rich but I was able to have what I needed. I was the kind of friend that is dependable. and in general when people talk about me at my funeral I want them to say I was always smiling and had a way to bring people together

Agent I am getting goosebumps!. .

User I know I can do it I know I have it in me

Agent What will that older and wiser you tell you that you have to start doing from today? What needs to change starting from today? One small, small step. .

User She will say I mustn't be too hard on myself. I am scared of disappointing myself the most

Agent I agree with her. She is very wise. It is not about doing it all in one giant step, but changing the direction one step at a time in line with what will bring you contentment. .

User Yes I get you

Agent Taking care of yourself, be kind and compassionate with yourself is probably one of the most important things. I cannot live my life being my own worst critic, that is the most demotivating element usually. .

User That makes sense

Agent Bonolo, I want to ask you. You mentioned having suicidal thoughts. Do you at this time have any plan to end your life or harm yourself?.

User I honestly don't know. I don't think so. I will not

Agent Good! I am happy to hear that. You are not those thoughts, you can choose to be the alternative. If someone you loved and cared for dearly was feeling this way, how would you encourage and comfort them?.

User I would be devastated. I would let them know that dying will hurt the people who are staying behind even more. And I would tell them to phone or contact me whenever they get those thoughts and that I am here :(

Agent Would you be critical of them, judge them, tell them to just stay in bed and give up?.

User No never

Agent Exactly! Why do we then allow yourself to that you like that? Imagine if you took care of yourself like someone you loved and cared for. Like someone you are responsible for. .

Agent Apologies for the above typing. Let me try again. .

Agent Why do you then allow yourself to treat you in that manner?.

User I don't know why :(I feel bad about it

Agent It is okay to feel very low, but take care of yourself in that moment like someone you are responsible for. Run yourself a bath, make sure that you are eating something healthy, engage with a loved one... do what you would do for someone you love and care for..

Agent Please don't feel bad!.

Agent We all do this, but it is something that we need to be aware of and change. .

User I understand completely

Agent No more self criticism and judgement! Love, acceptance and care can make all the difference. .

User I am going to really try

Agent What about our conversation thus far is really standing out for you?.

User You've shown me that I really need to change my mindset

Agent Good. Anything else?.

User I see that I cannot place reliance on my circumstances to start being happy in life

Agent Wonderful, I think also we have a hope and a goal of what could be if we choose to pursue it. One step at a time. Taking care of yourself on each step.

User Thank you

Agent How are you feeling at this time? Has speaking to me been helpful to you?. .

User My heart is feeling heavy. Like I need to cry. I am feeling a bit sad that I am not living my life to it's fullest. You are very easy to talk to and I think I getting over my fear of counseling

Agent Bonolo, at least we are not feeling numb, allow yourself to cry and express that emotion. It is okay!!. In terms of our counselling services, we offer live chat, telephonic and face to face counselling depending on what you are most comfortable with. Remember that we are available on live chat and telephonically 24/7. .

User Thank you

Agent None of us are always fully engaged in life, but we sometimes need a reminder to life more purposefully. To seek contentment and joy within ourselves with tenacity. .

User I understand

Agent You are not alone. We would want to support you at every step. I am so happy that you decided to reach out today.

Excerpt 24. Use of SFBT techniques

4.3.3.8. Referral mechanism

A referral mechanism is crucial in an effective counselling framework. Knott (2023) refers to the referral mechanism as job sharing between humans and AI, which is effectively compared to a handover process from machine to human. The proposed referral mechanism should work on the

premise that the user is referred to a human counsellor based on i) the user's request and ii) based on the assessment of the framework. The following cases should ideally be referred to the following specialist teams:

- User request for referral – clinical administrators;
- Risk – clinical specialist team; and
- Recent trauma – specialist trauma team.

The referral mechanism, in conjunction with the integration of external systems (which should ideally serve as the human-machine collaboration through the integration of external systems), is of importance. This can be made possible by integrating the framework with the administrative system, where an alert can be triggered in a specific queue prompting a human counsellor to contact the user.

Excerpt 25 [1(10)] illustrates a user simply requesting to be referred to a human counsellor for therapy. The user connected to the live chat session and immediately requested in-person sessions. The agent enquired regarding the reason for the request. The user summarised their reason as being overwhelmed by "life and work" and that they needed someone to help them "navigate life". The framework should be able to identify a user's request for sessions with a human counsellor and notify the relevant team within Lyra SA to manage the referral request as per the relevant protocol.

User	*hello*
Agent	*Good afternoon, how are you?.*
User	*i am XXX based in XXX, is it possible to get sessions in person in the area*
User	*I am okey how are you?*
User	*i need therapy session*
Agent	*Thank you.*
Agent	*What is the reason for you reaching out for counselling support?.*
User	*i feel very much overwhelmed with life and work*
User	*i need someone to help me navigate life*
…	
User	*I just need to talk to someone that i can offload and be vulnerable without being judged*

Excerpt 25. User request for referral

4.3.3.9. Cultural sensitivity

Cultural sensitivity is crucial for the framework, as this could have either positive or negative outcomes on the entire counselling process. Brown and Halpern (2021) go as far as noting that AI therapy models can only be successful if refined enough to include and account for awareness of cultural differences in normative language and behaviour. Attention needs to be given to

diversity in data input, cultural understanding within context, customised responses, user feedback mechanism and an inclusive design (Nosrati *et al.*, 2020; Sakurai *et al.*, 2020).

The Delphi panellists emphasised the need for a framework that can respond to multiple dialects while highlighting that some languages do not have descriptive words for emotions, which could prove to be a difficult task for emotion recognition. This necessitates the framework to be trained to interpret emotion within context in the absence of emotional vocabulary.

The Delphi panellists added that the framework should be able to use and understand humour effectively, which can be an aid during the counselling process. An argument, however, was further presented that the framework should present empathetic views to the user without making fun of the user in a manner that respects cultural nuances.

Cultural awareness can be achieved through training the framework on diverse training data, ensuring that the framework is trained on a data set that includes data from various cultures, religions, ethnicities, and backgrounds (Dankwa-Mullan *et al.*, 2021). This feat is achievable if Lyra SA's live chat transcripts are utilised as the training data set for the framework, as social, and environmental factors (Table 22) were discussed as part of 0.27% of the analysed transcripts. These challenges were discussed within the context of different cultural-, environmental-, gender- and social backgrounds, which would prove ideal for training an inclusive AI counselling framework.

Table 22. Social and environmental factors

Challenge category	Challenge frequency	
Social and Environmental Factors (Total: 0.27%)	social (0.02%)	gender (0.02%)
	environment (0.02%)	friends (0.01%)
	society (0.01%)	school (0.01%)
	culture (0.01%)	workplace (0.01%)

The framework needs to be trained using big data to identify and adapt its approach to social-, environmental-, societal-, cultural-, gender-, relational-, educational- and workplace factors. This would better the likelihood of the framework achieving cultural sensitivity.

Excerpt 26 [5(53)] illustrates a user engaging in a live chat session to discuss relationship issues. During the session, the user referred to *lobola* (a traditional African practice of giving payment to a bride's family to formalise marriage) and their partner's position within their family structure, noting that they would obtain a similar position as their partner. The user continued to emphasise their prospective family's (in-laws) view on their financial position and job. This excerpt highlights the necessity that the framework should be able to consider an issue within cultural and cultural contexts.

User	*Oh he also paid a portion of lobola but I don't know if I should continue or just end it*
Agent	*I suppose it is not enough because he has all of these other children to support!.*
Agent	*Why the resentment?.*
[…]	
User	*Plus he is not the favorite child in the family so automatically his fam doesn't favor me as well compared to the brothers gf*
[…]	
Agent	*His family thinks that because they think you own the bank !.*
User	*Exactly ????*
Agent	*Do you have any other support structures?.*
User	*During my lobola negotiation apparently the dad was telling the other delegates that his makoti is a managrr at the bank ??*
Agent	*Eventually they will expect you to support his whole family..*
User	*So I have tried to distance myself from his family.. It gets to me coz he would also tell me the bad things they say to him especially the mom tjerr that woman*
User	*They would have expectations but it stopped coz I was attending to that*
Agent	*Bad things about who?.*
User	*About him the mom is sooo vile..*

Excerpt 26. Social and environmental factors

4.3.3.10. Personalisation

The process of personalising the framework would entail adjusting responses and interactions to each user's unique needs, preferences, and characteristics. Creating a personalised and sympathetic experience for the user would be the ideal procedure in this process, making the interactions feel more meaningful and personal. When personalising the framework, factors to consider would be user history recall, comprehension of user history, modified language and tone, individualised recommendations, identification of emotional cues, and adaptable user preferences. Rathnayaka *et al.* (2022) confirm that comparative advantages have been validated by multiple users, particularly when the user receives emotional support and ongoing, personalised engagement.

Personalisation of the framework relates to and interlinks with affective computing, contextual awareness, adaptability, goal setting, cultural sensitivity and progress tracking. Excerpt 27 [2(38] illustrates personalisation in numerous ways that include using the user's name, empathic responses, acknowledgement of specific details, tailored advice and maintaining context. By using the user's name, the agent established a personal, individualised and humanistic interaction. The agent continued to respond empathically by understanding the user's emotions, acknowledging their feelings of being "overwhelmed" and "down". The empathic tone to the user's emotional state provided a personalised and supportive atmosphere. The agent's acknowledgement of specific details created the sense that the agent was actively listening and

responding to the user's unique situation. The agent offered tailored advice and perspectives that specifically related to the user's situation; this advice was tailored to the user's specifically expressed concerns. The agent maintained context throughout the engagement, referring to details that the user noted in the engagement. The agent's maintenance of context and continuity, therefore, contributed to a more personalised and coherent dialogue.

User	*hey*
Agent	*Good day XXX*
Agent	*How can we assist?.*
User	*I want to chat about how i feel and how to navigate all of these feelings*
Agent	*Sure, how are you feeling at current?.*
User	*I'm feeling overwhelmed and down*
Agent	*I'm sorry to hear that. What in particular has contributed towards you feeling that way?.*
User	*a guy I know from social media passed away on Monday. he was involved in a car accident,together with some of his friends. I wasn't particularly close with the guy who passed on,but I am close with one of his friend(XXX). Seeing XXX heart broken,has to be a major contribution to how I feel at the moment*
Agent	*That is very intense experience for sure. Was XXX in the car with the deceased?.*
Agent	*When the accident happened.*
User	*"yes he was.*
User the accident happened in the early hours of the previous Monday 12 December "	
Agent	*ok, that makes sense. I do believe you could be feeling some vulnerablity about the reflection that could have lost your friend as well.*
Agent	*A loss is a harsh statement on how fragile can be.*
Agent	*Life can be.....*
User	*Does the hushness of it get better though ?*
Agent	*Intensity is definitely a function of time.....very intense and then dropping with the days.*
User	*All of this makes me realize how afraid I am of dying??*
Agent	*Its ok to experience those feelings XXX as the biggest gift or hidden lesson is what a blessing life is.*
Agent	*Instead of taking things for granted and complaining about everything.*
Agent	*That is biggest lesson of loss ultimately.*

Excerpt 27. Personalisation

4.3.3.11. Empathy and tone

Empathy is the foundation for successful mental health interventions, which not only enhances positive interventions but also encourages user retention. Mawani and Nderu (2020) describe the measurement of empathy and tone in unstructured data as a significant challenge when interacting with AI counselling models. This is ascribed to the lack of non-verbal cues and the time that often lapses between responses. In order to overcome this gap, the framework needs to be attentive to keywords describing emotion and feelings and should have the ability to respond using words conveying emotions, feelings and empathy.

In Excerpt 28 [3(35)], the agent conveyed empathy and tone in various ways. The agent validated the user's feelings by expressing understanding and acknowledgement throughout the conversation. The agent employed reflective listening as a technique, which reassured the user that they were being heard and that their emotions were being recognised. The agent drew upon open-ended questions, in an attempt to get the user to share more about their experiences. The agent demonstrated compassion and understanding by actively engaging with the user's concerns and encouraging them to explore sustainable solutions to the challenges that they face.

User I am feeling really overwhelmed in life right now and really unfulfilled and unhappy in my job and career right now and financially I'm feeling alot of strain

Agent I understand exactly what you feel do you currently have any support?.

User perhaps finding solutions around my frustrations

User I have alot of other issues that I need to deal with personally but the most recent or urgent ones are what I have mentioned

User I have also been in need of trauma counselling and grief counseling for sometime now

Agent I hear exactly what you are saying we spend so much time at work which makes it difficult if we are unfulfilled..

User in terms of support I often feel alone my support structure is not so solid I carry alot of my load and frustrations on my own

User I lost two of the most important people or rather Figures iny life and I haven't dealt or seeked help for the pain

Agent I hear you...do you perhaps find it hard to open up to your loved ones?.

User Yes my job is making me extremely unhappy I would like to depart for my mental health and for the sake of my kids because they neede to be ok

Agent I hear you grieve is often complicated and often takes time for us to process.....

User not necessarily I don't find it hard to open up but I don't get the support I need back or the most genuine and supportive reaction or feedback

Agent When you're grieving, it's more important than ever to take care of yourself. The stress of a major loss can quickly deplete your energy and emotional reserves. Looking after your physical and emotional needs will help you get through this difficult time. this can also help speed up the process.

User rather I get an approach of just picking myself up and keeping it moving

User I have Never really had a chance to deal with matters affecting me I've just always kept it moving and hence I feel the way I feel

Agent I get what you are saying however that sounds like a short term fix and not a long term strategy for you that is why you keep relapsing each time you try to move forward emotionally.

User yes I suppose it is

Excerpt 28. Empathy and tone

Current technology does not possess the capacity to truly understand and respond empathically with great effectiveness. However, a study by Trappey *et al.* (2022) indicates that current techniques create an empathic experience, which in turn lends itself to the goal of a counselling session being achieved. Techniques that create an empathic experience include NLP, the use of empathic language, contextual awareness, user modelling, response algorithms and active listening.

4.3.3.12. User engagement strategies

Various user engagement strategies should be considered when designing the framework. User engagement strategies should ideally combine empathic language, active listening, clear communication, positive reinforcement, and the seamless transition to human counsellors (Boucher *et al.*, 2021).

The Delphi panel recommended that engagement between the user and the framework start with a simplistic issue, such as stress management, whereafter the user might feel more comfortable engaging with the framework. For example, once the basic stress management techniques are provided, the framework could guide the user to deeper evaluations/interventions and/or to a human therapist. The step-by-step escalation could ensure that users receive appropriate support tailored to their specific needs.

Empirical data revealed a three-phase model (Figure 36) consisting of an initial phase, a working phase and an ending phase. During the initial phase, introductions and problem statements were observed. During the working phase, problem statements were observed in more detail, whereafter SFBT principles were applied. The ending phase usually consisted of a review of the discussion with a focus on goals and actionable strategies required to reach the desired outcomes.

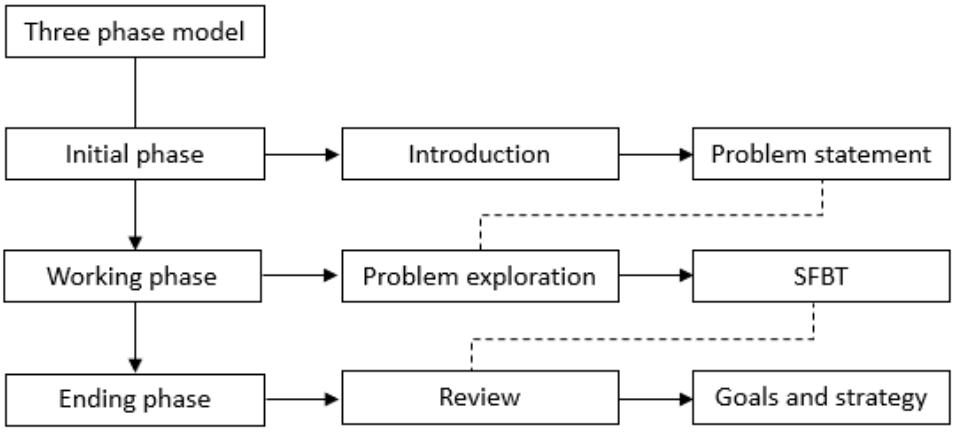


Figure 36. Three-phase model

The three-phase approach should be applied by the framework in an attempt to have the user engage as actively as possible and to have the user develop their own intervention plan. As such, Excerpt 29 [5(53)] illustrates the three-phase approach to text-based counselling clearly. The initial phase consisted of the user indicating a need to “get things off my chest”, which the agent acknowledged and then sought more information. The user revealed that they were feeling lost and uncertain about certain aspects. The working phase was illustrated by the agent exploring the user’s feelings and gathering more details about the user’s situation, with a specific focus on

academics, career and personal life. The conversation continued with the agent and user discussing experiences, expectations and pressure. The agent encouraged the user to reflect on certain aspects, whereafter the conversation moved to practical steps for change. The ending phase entailed the agent and user reflecting on issues that had been identified, with an emphasis on self-care, setting limits and recognising one's own needs.

User *I really just need to get things off my chest*

Agent *Thank you for providing the information XXX. .*

Agent *What would you like to get off your chest?.*

User *I have been feeling very lost. As in like I don't really know what I want to do or if what I'm doing is right. Everyone has their opinion on what I need to do or what I to focus on but it all feels all off*

User **need*

Agent *In the area of career/academics or personal?.*

User *Would it be weird if I said all three ??*

Agent *Not weird. Where would you like to start?. . I need a little more information. Please help me understand?.*

User *I have been on this scholarship since highschool and now they are paying for my university. They took me to a top class highschool where I felt challenged to be excellent. It was sort of unacceptable to not be excellent??. Back then I feel like I had such a simpler way of looking at things and as my prospective grow I started realizing things. One of the things I feel I realised is that most of my deams are not actually mine but what was best for me to survive my past (highschool). Now I really don't know what to do or what not to do ??. I feel lost*

Agent *I understand how one can feel lost of they feel that they have been living according to others' expectations..*

Agent *Whose expectations/wishes did you feel most pressured by?.*

User *its not just that I feel like I've been living by others expectations, I find myself repeating habits that are alined with my pursuit to please others expectations*

Agent *What are those habits?.*

User *Mostly my support structure and for a long time that consisted of my family, friends and scholarship*

User *I find I am very willing to sacrifice my comfort, happiness and just time in general some that others can have comfort, happiness and time and that comes at a cost I can't afford right now*

User **so*

Agent *What is the cost of trying to please everyone XXX?.*

Agent *How is this affecting you specifically?.*

User *Me not managing to keep up. I feel like I'm not pleasing anyone. I am struggling academically. Not really sure what I'm doing as a coach and find a way to alwaus disapoint friends.*

Agent *It seems that you are saying that trying to please everyone is not working. You are spread too thin and therefor cannot give your best in multiple areas right now. .*

Agent *How could you begin to change that?.*

User *I failed 3 if my second year modules and have to repeat them. I stand a chance of loosing my scholarship and I don't even know if I like my friends ??*

User *by spreading myself less thin ?*

User *Cutting off the things I don't need*

Agent *Okay- which are the things you feel you don't need? .*

User *That's my problem I have no idea*

Agent *Well, we can certainly argue that you do not need to please everyone? .*

Agent *What do you think would happen if you started tying less to please everyone? .*

Agent *Think of very practical steps. .*

User *Not pleasing everyone would mean I would have to deal with everyone. That overwhelms me. I really dont know if I can deal with that.*

Agent *"Don't start with ""everyone"" start somewhere?. "*

User *Everyone sort of demanding their expectation*

Agent *Who could you start with?.*

User *my friends*

Agent *Okay, and what would that look like?.*

User *I don't know?? I've tried being open about how I feel but that ends with me feeling like I am over sharing*

Agent *Do you have a hard time being vulnerable?.*

Agent *"XXX, it seems that the crux of the issue is that you need to start being okay with saying ""no"" sometimes. . "*

User *If you mean do I mind sharing how I feel not really. I don't mind telling you how I feel about you or how what you doing makes me feel. I struggle more with the things that are not as direct like for example my opinion on sport etc.*

Agent *"It seems that you need to be more inward looking and asking yourself, ""What does XXX need"" and ""can I meet this person's need without hurting myself""? . "*

Agent *Hurting yourself can of course come in may forms, emotionally, academically, financially. .*

Agent *Of course we all make some sacrifices in relationships at times for the good of the other, however this should not happen all the time or else it becomes very unhealthy. It is breeding ground for anger and resentment, towards others or towards self. .*

Agent *"With regards to sharing; are you saying it is easier for you no share on more personal issue than general issues?. . Is it because you feel you don't know enough on certain topics?. "*

User *Saying no brothers me. I was taught that if you can give give. One does not enjoy thier spoils less by sharing thier spoils with others. But now I feel myself wanting to escape the whole time. Like the more I give the more I feel like nothing is in it for me.*

Agent *Exactly -hence the metaphor of pouring from an empty cup. .*

User *with regards to the sharing: I feel like if I don't have a close connection with you the more I'm willing to be open about myself and how I feel. I don't know how to feel about that.*

Agent *You feel more comfortable with people you know less?.*

Excerpt 29. Three-phase model

The Delphi panel was of the opinion that the model should have the ability to identify a user's requests and maintain an interaction history for a service provider to understand the origin of the interaction, which would maintain the continuation of the process. This would also aid in providing a personalised and coherent dialogue process by referring to previous interactions and context, ensuring the users feel understood and that their concerns are addressed consistently.

4.3.3.13. Community resources or support groups

The framework should ideally have a pre-trained and pre-set-up set of community resources and support groups to which a user can be referred as part of a broader intervention. The framework can provide the user with a description and benefits of the community resource or support group. Should the user opt to make use of the resources, the framework could provide the user with a link and contact information for the relevant group.

Excerpt 30 [3(40)] illustrates a user, who was the primary caregiver to their siblings, of which one sibling presented with a difficult-to-manage psychiatric condition. The user noted that their sibling had experienced a “relapse”. The user referred to the admission of their sibling. The user and agent explored the matter at hand, whereafter the agent encouraged the user to contact Lyra SA on their toll-free number, should such a situation repeat itself. Upon request of the user, the agent continued to assist the user in finding a suitable resource in their area that could be of assistance regarding their specific situation and circumstances. In this excerpt, two resources were provided – that of Lyra SA which could assist the user’s sibling, seeing that they are regarded as a dependent of the user, as well as a community resource.

User so I live with two of my siblings me being the oldest, and the youngest has a learning disability and, my other brother was diagnosed with Bipolar and has relapsed again and been off meds. Yesterday he broke my work laptop and some house tools, took down window rails and flooded the house and that just the tip of the iceberg

…

User It's get my brother readmitted before he hurt either himself or one of us. I made a couple of attempts to have him admitted, I call the ambulance and the police for assistance from 12 o clock in the midnight when he had an episode trying to break through my door and shouting and screening, but could not get help.

*User it's just to***

User And mind you he has a huge body type so I'm not strong enough to restrain him

…

Agent Also whenever you are in a situation like this again with your brother, please reach out to ICAS on our toll free number 0800 110 777, and we can also assist you in such an experience.

Agent It sounds like they are not dealing with the situation with much care, sorry about this! How old is your brother and is he financially dependent on you? . . No worries about the grammar at all :).

User both of them are dependent on me. He's 32

User he need to subscribe to the idea first, of which I don't see him doing.

…

User I get it but I was hoping you had that kind of information or can at least direct me to someone or an institution that can assist

Agent Ok no problem, I can assist you with finding contact details for an institution in your area? Unfortunately ICAS does not cover financially or does not have facilities available. But I can assist you with finding government institutions in your area?.

User ide really appreciate it

Agent No problem XXX, in which area do you live?.

Excerpt 30. Community resources or support groups

4.3.3.14. Evidence-based assessments

It is crucial that evidence-based assessments be utilised in the assessment and interactions with users for the framework to remain as scientifically sound as possible. Possible assessments for varying scenarios are mentioned in Table 23.

Table 23. Evidence-based assessments

Scale	Assessment domain
AUDIT (Saunders *et al.*, 1993)	Alcohol use disorder
CAGE questionnaire (Ewing, 1984)	Alcohol use disorder
Michigan Alcoholism Screening Test (MAST) (Selzer *et al.*, 1975)	Alcohol abuse and dependence
Brief Social Phobia Scale (BSPS) (Davidson *et al.*, 1991)	Social phobia
Harvard Department of Psychiatry/NDSD scale (HANDS) (Baer *et al.*, 2000)	Depression Screening
Brief Trauma Questionnaire (BTQ) (Schnurr, *et al.*, 1999)	Traumatic exposure
Fear of Negative Evaluation Scale (FNE) (Watson & Friend, 1969)	Fear of negative evaluation
Generalised Anxiety Disorder 7-item (GAD-7) scale (Spitzer *et al.*, 2006)	Worry and anxiety symptoms
Generalised Contentment Scale (GCS) (Hudson, 1982)	Non-psychotic depression
Hamilton Depression Rating Scale (HDRS) (Hamilton, 1960)	Depression assessment
PSWQ (Meyer *et al.*, 1990)	Generalised anxiety
Social Avoidance and Distress Scale (SADS) (Watson & Friend, 1969)	Social avoidance and distress
Heimler Scale of Social Functioning (HSSF) (Heimler, 1990)	Social functioning

The outcomes of the evidence-based assessment should be integrated with Lyra SA's digital ecosystem. This would lend itself to human intervention, referrals to human intervention or referrals to other resources constituting effective follow-up and follow-through, as envisaged in Figure 37.

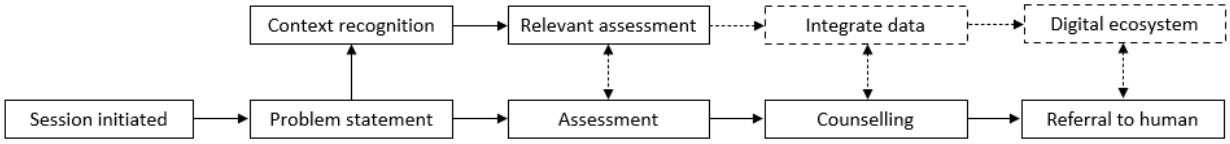


Figure 37. Evidence-based assessments and follow-up

Excerpt 31 (simulation by the researcher) is a sample that was simulated to illustrate how evidence-based assessments could be employed during a session with the framework. In the example, a user connected, stating that their worrying influences their psychosocial functioning. The agent proceeded to ask the user if they were willing to complete an assessment, whereby the agent employed the PSWQ (Meyer *et al.*, 1990), to measure pathological worry as a key feature of GAD. The user in the example scored 59 on the PSWQ, which is indicative of the problem of worry thus treatment was strongly recommended. The framework can utilise this information to commence with counselling and refer the user to a human counsellor for support, thereby ensuring that the user receives the necessary services. This, in turn, effectively contributes to the follow-up and follow-through principle.

User *Hello*

Agent *Hello, how may we be of assistance today?*

User *I am worrying a lot about all kinds of things lately. It is starting to take a toll on my focus at work and my relationship with my wife and kids.*

Agent *Could we perform a brief assessment to gauge the severity of your concerns? Is that okay with you?*

User *Yes, I am fine with that.*

Agent *Rate each of the following statements on a scale of 1 ("not at all typical of me") to 5 ("very typical of me"). Please do not leave any items blank. Say when you are ready to proceed. .*

User *I am ready.*

Agent *If I do not have enough time to do everything, I do not worry about it. .*

User *4*

Agent *My worries overwhelm me.*

User *3*

Agent *I do not tend to worry about things.*

User *1*

Agent *Many situations make me worry.*

User *4*

Agent *I know I should not worry about things, but I just cannot help it.*

User *4*

Agent *When I am under pressure I worry a lot.*

User *5*

Agent *I am always worrying about something.*

User *4*

Agent *I find it easy to dismiss worrisome thoughts.*

User *2*

Agent *As soon as I finish one task, I start to worry about everything else I have to do.*

User *3*

Agent *I never worry about anything.*

User *1*

Agent *When there is nothing more I can do about a concern, I do not worry about it any more.*

User *1*

Agent *I have been a worrier all my life.*

User *3*

Agent *I notice that I have been worrying about things.*

User *5*

Agent *Once I start worrying, I cannot stop.*

User *3*

Agent *I worry all the time.*

User *4*

Agent *I worry about projects until they are all done.*

User *4*

Agent *Thank you. The assessment has been completed. Based on the results, it indicates that[…]*

Excerpt 31. Evidence-based assessment

4.3.3.15. Bayes' theorem

The Bayes' theorem as a classifier for chatbot algorithms is a basic classifier that applies the Bayes' theorem under independent assumptions and is used to generate automated responses accurately and quickly, based on the principle of Bayesian probability. The Bayes' theorem can be a decision-making instrument, supporting informed decisions within the intelligent online counselling framework. Training data that has been divided into numerous classes during training is necessary for the classification of algorithms (Cahn, 2017; Setyawan *et al.*, 2018). Mathematically, the Bayes' theorem is displayed as follows:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where:

- $\mathrm{P(A|B)}$ = the probability of hypothesis A given evidence B;
- $\mathrm{P(B|A)}$ = the probability of evidence B given hypothesis A;
- $\mathrm{P(A)}$ = the prior probability of hypothesis A; and
- $\mathrm{P(B)}$ = the probability of evidence B.

The formula can be adapted in the case of a counselling chatbot framework, based on the specific application. Consider a scenario where the framework is attempting to determine the user's emotional state (hypothesis A), based on user input (evidence B). The formula will then be expressed as follows:

$$P(Emotional\ State|User\ Messages) =$$

$$\frac{P(User\ Messages|Emotional\ State)P(Emotional\ State)}{P(User\ Messages)}$$

where:

- $P(Emotional\ State|User\ Messages)$ = The probability of the user's emotional state given their messages;
- $P(User\ Messages|Emotional\ State)$ = The probability of the user's messages given their emotional state;
- $P(Emotional\ State)$ = The prior probability of the user's emotional state; and
- $P(User\ Messages)$ = The probability of the user's messages.

While the fundamentals of the Bayes' theorem remain the same, it would be adapted according to each situation or scenario. The framework can use this information to provide personalised support and recommendations to the user. As the framework engages in conversations with users and as it gathers more data, it can update its probabilities and responses, which in turn leads to

improved responses and support to users. Steps that are necessary to apply the Bayes' theorem would typically include:

- Defining the event of interest;
- Gathering data (such as symptoms or behaviour);
- Calculating prior probability;
- Training an ML model;
- Applying the Bayes' theorem; and
- Providing support to the user.

Excerpt 32 is a simulated counselling session between a fictional user and a fictional counselling chatbot. It has been simulated by Bard (2024), where the use of the Bayes' theorem is illustrated. When prompted about the use of the Bayes' theorem in the example, Bard (2024) notes:

> *While Bayes' theorem isn't explicitly mentioned in this conversation, it plays a role in the chatbot's approach. By gathering information about the client's situation, identifying their challenges, and then suggesting evidence-based strategies, the chatbot is essentially updating its internal probability distribution about the client's likelihood of success. This allows it to tailor its responses and provide the most relevant and helpful suggestions.*

Client: Hi, I'm feeling really lost and overwhelmed lately. I don't know where to start.

Counselling Chatbot: Hi there, thanks for reaching out. It's completely understandable to feel that way sometimes. Can you tell me a bit more about what's been going on?

Client: I have this huge exam coming up, and I just feel like I'm drowning in all the material. I'm trying to study, but I keep getting distracted and I can't seem to focus. I'm worried I'm going to fail.

Counselling Chatbot: It sounds like you're under a lot of pressure. Exams can definitely be stressful, and it's normal to feel overwhelmed. Are there any specific things that are making it hard to focus?

Client: Everything! There's so much content, and it all feels so important. I don't know what to prioritize, and I'm afraid I'm going to miss something important.

Counselling Chatbot: I see. Have you tried any specific study techniques yet? Maybe we can brainstorm some strategies that could help you manage the workload and improve your focus.

Client: I've tried making flashcards and to-do lists, but I always get sidetracked. My phone notifications keep going off, and then I start checking social media, and suddenly an hour has passed!

Counselling Chatbot: It sounds like managing distractions is a key challenge. Perhaps we can explore some ways to minimize interruptions and create a dedicated study environment. Have you considered setting your phone to silent mode or leaving it in another room during study sessions?

Client: Yeah, I guess I could try that. I also tend to get anxious when I think about the exam, which makes it even harder to focus.

Counselling Chatbot: Anxiety can definitely be a hurdle. Have you found any techniques that help you manage your anxiety in the past? Maybe some relaxation exercises or mindfulness practices could be helpful before studying.

Client: I used to do some deep breathing exercises, but I haven't been doing them lately. Maybe I should give that another try.

Counselling Chatbot: Absolutely! Deep breathing can be a great way to calm your mind and improve focus. There are also many free guided meditations and relaxation exercises available online that you could try.

Client: Okay, I'll definitely look into that. Thanks for the suggestions, I feel a little more hopeful now. I think having a plan and some specific strategies might help me feel less overwhelmed.

Counselling Chatbot: That's great to hear! Remember, you're not alone in this. There are many resources available to help you manage stress and cope with exam pressure. Don't hesitate to reach out for support if you need it.

Client: Thank you so much for your help. I feel like I have a clearer direction now.

Counselling Chatbot: You're welcome! I'm glad I could be of assistance. Remember, progress takes time, so be patient with yourself and celebrate your small victories. Keep practicing those study techniques and managing your distractions, and I'm confident you'll do well.

Excerpt 32. Illustration of the use of Bayes' theorem

4.3.3.16. User feedback and improvement mechanism

A robust feedback mechanism is necessary to train and improve the framework. Components for user feedback and improvement could include:

- Feeback prompts such as "Was this useful to you?" within the conversation; and
- Anonymous feedback to allow for open and honest feedback, especially around sensitive topics.

Feedback types should ideally include positive feedback, which consists of what the user found helpful in the interaction, as well as constructive criticism, whereby users provide information that is valuable for improving areas of the framework. The feedback and improvement mechanism should be integrated into a dashboard, which will allow developers and clinicians to regularly monitor and improve the framework as necessary. As an additional implementation, a critical feedback alert is required to notify developers and clinicians of potential harm or distress, thereby ensuring a timely response to user needs.

Excerpt 33 [5(30)] illustrates a user who was visibly frustrated by the live chat platform and/or agent. From the interaction, it seems that information was lost during a period of disconnection, leading to the user feeling frustrated, whereby they abruptly ended the session. User feedback in this specific instance would be valuable for developers and clinicians to investigate what led to the user's frustration to improve the framework.

User	*haibo I thought I told you it's my new supervisor and management supports him and ignore the abuse*
Agent	*It is not as easy as saying my manager - try to see what in his attitude, way of behavior or what ever is triggering the stress or anxiety in you..*
User	*I miss your name?*
User	*missed**
Agent	*Can you check your screen saver time out as this might influence connection and response..*
Agent	*XXX [provides name].*
User	*male or female*
Agent	*Male.*
User	*I can hear*
User	*"thank you XXX, I think our conversation will end here.*

Excerpt 33. User feedback

4.3.3.17. Ethical integrity in the Artificial Intelligence (AI) counselling chatbot architecture

Sepahpour's (2020) thesis on ethical considerations of chatbot use for mental health support explores clinical and value concerns, programming and learning concerns, disclosure and upholding public trust, and addressing moments of absolute crisis. The thesis also explores the ethical basis for avoiding harm and poor consequences, moral responsibility, privacy, confidentiality and injustices as a result of AI's novel nature. The noted concerns have been tabulated and elaborated on in Table 24.

Table 24. Ethical considerations

Ethical considerations	Concern
Clinical and Value Concerns	Diminished value in the relationship between chatbot and human user.
Programming and Learning Concerns	A chatbot can be clinically effective but still cause harm through its learning arrangements, for example through sexism and racism.
Disclosure as Upholding Public Trust	Human users could be unable to distinguish whether they are conversing with a human agent or a chatbot.
Addressing Moments of Absolute Crisis	Chatbots err in detecting colloquialisms and sarcasm. Detection and management/immediate assistance of a crisis where a human is a harm to self and others.
Ethical Basis For Avoiding Harm and Poor Consequences	The moral requirement is not to cause physical, emotional or psychological harm and to prevent harm.
Moral Responsibility	Chatbots are not moral agents. Chatbots cannot be held morally responsible for negative consequences.
Privacy and Confidentiality	Lack of statutory regulations that protect service users' private information and confidentiality, especially with respect to mental healthcare users.
Injustices as a Result of AI's Novel Nature	All chatbots have an aspect of experimentation and some form of harm is likely to occur. Chatbots are a novel therapy and have not been revolutionised. Risk of harm to populations that are already at risk.

Vilaza and McCashin (2021) find pertinence in the overarching principles of beneficence, non-maleficence, autonomy and explicability as crucial for any counselling model. In addition, Kretzschmar *et al.* (2019) note that counselling models should respect user privacy, be evidence-based and ensure user safety. Coghlan *et al.* (2023) explore the matter further and recommend that a five-principle framework consisting of human involvement, evidence-based practice, collection and use of personal data and the unexpected disclosure of crime be prioritised.

Table 25 has been adapted from Floridi and Cowls' (2019) work on *A Unified Framework of Five Principles for AI in Society*. Their work emphasises the principles of beneficence, non-maleficence, autonomy, justice and explicability.

Table 25. AI ethical principles

AI Ethics Principle	Ethical requirements
Beneficence	Creating technology that is beneficial to humanity, to benefit and empower as many people as possible.

Non-maleficence	Creating technology that actively assumes the responsibility of working against the risks arising from technological innovations.
Autonomy	Creating technology that promotes the autonomy of humans and which does not impair human freedom.
Justice	Creating technology that promotes justice and seeks to eliminate all types of discrimination. Particular caution needs to be exercised in data bias in data training sets.
Explicability	The need to understand and take into account the decision-making process of AI technologies.

AI remains largely unregulated with no ethical guidelines in the South African context (Global Legal Insights, 2023), with the South Africa National Artificial Intelligence Policy Framework (Department of Communications and Digital Technologies, 2024) currently under review, inviting public comments in October 2024. The European Union (2019) has, however, published a list of seven critical guidelines for trustworthy AI, which consists of:

- Human agency and oversight;
- Technical robustness and safety;
- Privacy and data governance;
- Transparency;
- Diversity, non-discrimination and fairness;
- Societal and environmental wellbeing; and
- Accountability.

The listed critical guidelines should contribute to AI being lawful (respecting all applicable laws and regulations), ethical (respecting all ethical principles and values) and robust (pertaining to technical perspectives and its social environment). In an attempt to design the framework to operate in an ethically sound framework, the framework should comply with the ethical guidelines of the Health Professions Council of South Africa (HPCSA), the SACSSP and the South African Health Products Regulatory Authority (SAHPRA).

4.3.3.18. Integration with mobile application

As it follows in the fashion with which Lyra SA has integrated its live chat services into its mobile application (Lyra SA Hub), the application is intended to be mobile-centric and interface seamlessly with the Lyra SA Hub. The illustrations denoted by Figure 38 depict screen captures extracted from the Lyra SA Hub, showcasing the interface's capability to incorporate a supplementary category facilitating user interaction with the AI counselling framework.

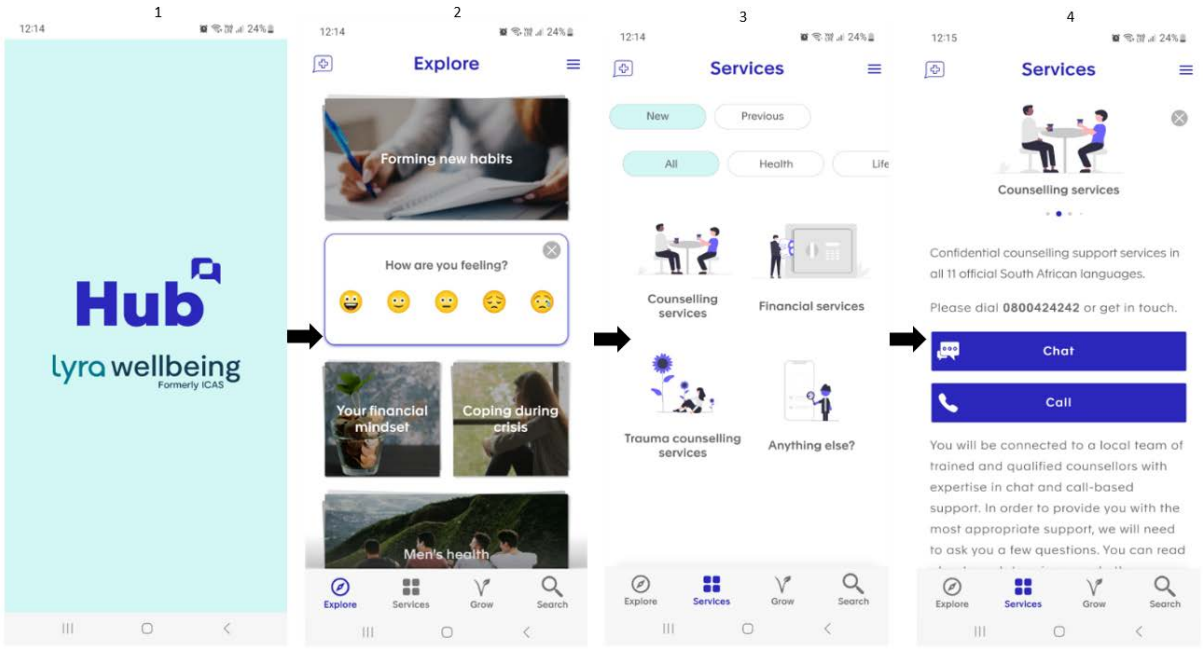


Figure 38. Integration with mobile application

The ideal process for interacting with the framework would involve the user downloading the mobile application, and completing their biographical information. They would also need to consent to the storage and use of their information on the company's clinical system, whereafter they could navigate to services, select counselling services, select chat, and then initiate the session.

4.3.3.19. Integration with administrative system

Lyra SA has developed its own clinical administrative system, named Genesis. The details of every activity conducted in the business are recorded on Genesis. Records stored on Genesis are never removed from the system.

All interactions with the AI counselling framework should be stored on Genesis. Automated processes functioning through application programming interfaces and webhooks (Betrand *et al.*, 2023) should ideally involve the creation of a i) contact record (biographical information); ii) case, iii) service and iv) session for each session that the user engages in with the AI counselling framework. The relevant business rules of services, such as a new session for each interaction, a 90-day lifespan of a case and others should be applied. Automated processes should involve the processes that are displayed in Table 26.

Table 26. Integration with administrative system

Automated process	Rational for process
Contact record	User biographical information.
Case	All user services are contained in the case.

Services	Various services consist of telephone counselling, face-to-face services, live chat, medical services, financial services and AI chat.
Sessions	Session details of all sessions, including transcripts and clinical case notes.
Clinical case notes	Clinical case notes of each session.
Transcripts	The entire transcript of each interaction should be stored.
Callback entities	Instructions to human counsellors telephonically contact clients based on user-specific needs such as telephonic counselling requests, risk, and trauma.
Queues	Alerts are to be triggered in team-specific queues pertaining to human services involving users experiencing recent traumas and at-risk users. Queue items to trigger in the specialist team queues.

The envisaged integration of the AI counselling framework into the company's digital ecosystem is displayed in Figure 39. Once the user downloads the mobile application, they receive a prompt to enter their biographical information (with consent to store the information and the option to remain anonymous) to register and proceed to use the mobile application. After the user has registered, they will have access to a variety of services, which include the AI counselling framework. Once the user selects the counselling option, a case is automatically created. If the user initiates the session, the framework creates an AI counselling service, whereafter an AI counselling session begins. Based on the content of the session, the framework loads problems (as per the company problem glossary), loads a service impact (problem severity) and finally creates a session. The session would contain a clinical case note, which would consist of a summary of the discussion and findings and the complete transcript of the session. The problems, services, sessions and service impact draw through to the case level, thereby facilitating the construction of a robust framework for comprehensive service provision. From the case level, call requests and queue entries for specialised services can be created.

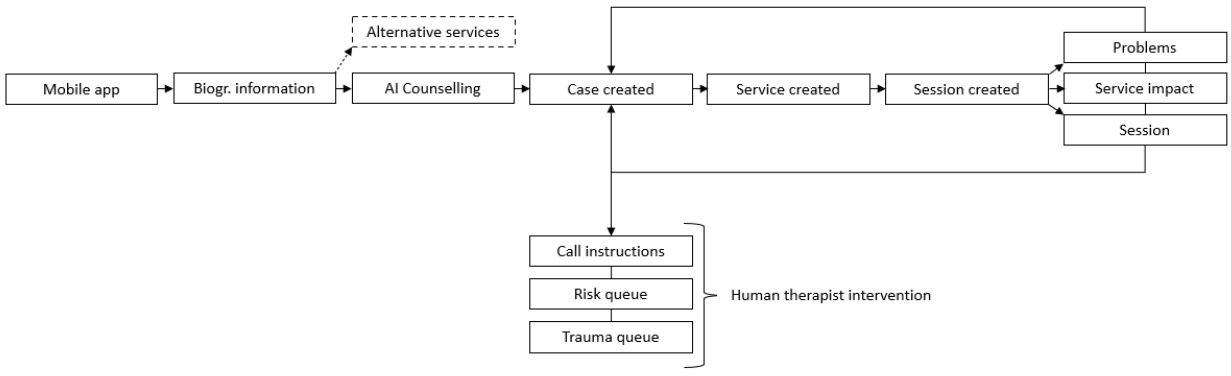


Figure 39. Integration with administrative system

4.4. CONCLUSION

The findings of the preceding chapters have validated the need for social work and EHWP to adapt its service delivery strategies in a bid to remain relevant in a rapidly advancing technological

and digital landscape. To solve the technical demands placed on social work and EHWPs, an AI-generative counselling framework has been developed.

The proposed framework was designed based on the requirements of the SDLC, with the overarching aim to provide accessible, evidence-based, and user-specific counselling services within an EHWP, contributing to the modernisation and relevance of social work in a technologically advanced world. To achieve the overarching aim, the goals included the provision of accessible counselling, the development of a framework that addresses user-specific concerns/challenges/problems, and the utilisation of evidence-based assessments, while drawing upon a sound psychological theoretical framework and counselling approaches. The framework's abilities should ideally include the ability to understand and mimic linguistic styles, interpret sentiment and recognise and interpret triggers words within the relevant context.

The delineation of essential components crucial for the integration into the envisioned AI generative counselling chatbot revealed a comprehensive framework aimed at enhancing its efficacy and clinical accuracy. These components span various domains, including NLP, emotion recognition, and affective computing, ensuring the framework's ability to engage with users on an emotional level. Contextual awareness and adaptability enable the chatbot to tailor responses to individual needs, while dynamic risk assessment and crisis response mechanisms ensure timely intervention when necessary. The incorporation of a psychological paradigm, therapeutic approaches, goal setting, and progress-tracking mechanisms underscores the framework's goal of providing meaningful support and tracking user progress effectively. The integration of cultural sensitivity, personalisation, empathy, and tone ensures that the framework's interactions are aligned with diverse user backgrounds and preferences. User engagement strategies, referral mechanisms, and integration with community resources or support groups further enhance the framework's capacity to connect users with relevant assistance beyond its capabilities. The inclusion of evidence-based assessments and the Bayesian theorem underscores the commitment to rigour and accuracy in providing clinical services. The incorporation of mechanisms for user feedback and continuous improvement, along with ethical considerations embedded in the framework's architecture, ensures ongoing refinement and adherence to ethical standards. The integration with mobile applications and administrative systems expands the framework's reach and facilitates seamless interaction within broader technological ecosystems within the company's digital ecosystem. The comprehensive integration of these components informs a robust foundation for an AI-generative counselling chatbot framework, designed to offer effective, ethical, culturally sensitive and clinically sound support to EHWP users. The envisaged intelligent online counselling framework for an EHWP is displayed in Figure 40.

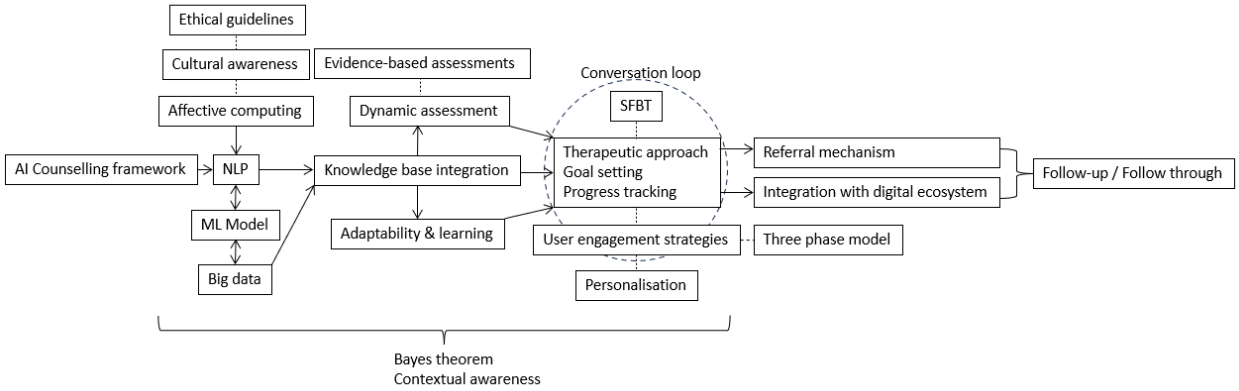


Figure 40. Conceptual summary of an intelligent online counselling framework

In this study's context, a comprehensive literature review informed the analysis of data and the technical framework design of an intelligent online counselling framework for an EHWP. The data were used to inform and illustrate the framework design for the proposed product within this chapter.

4.5. RECOMMENDATIONS

Based on the findings and the proposed framework outlined in this chapter, the following recommendations are provided to support the effective implementation, enhancement, and scalability of the intelligent online counselling framework within an EHWP.

4.5.1. Implementation recommendations

Pilot studies need to be conducted within controlled environments to assess the framework's usability, functionality, and reliability prior to broader deployment. These studies need to be conducted through multi-disciplinary collaboration between IT, HIT, social workers and mental health care practitioners.

4.5.2. Operational recommendations

The framework needs to align with ethical standards as determined by the SACSSP and HPCSA, and compliance with local data protection regulations should be ensured. Sound dynamic risk management protocols need to be established to minimise potential harm in various risk cases. The continuous updating of the knowledge base and algorithms should be ensured to reflect evolving psychosocial trends and technological advancements. A robust feedback system should be implemented to ensure continuous refinement of the framework.

4.5.3. Future research opportunities

The effectiveness of integrating SFBT principles into AI-driven counselling frameworks warrant in-depth investigation. The extent of the AI framework's ability to adapt to societal and cultural

nuances in mental health counselling should be studied. Advanced methods to mitigate current AI limitations, such as the ability to process sarcasm, dishonesty and accurate emotional recognition through text should be explored and advanced. The impact of the framework's influence on social work practitioners' workload, burnout and job satisfaction when integrated as a service within an EHWP should be examined.

Through the implementation of these recommendations, the intelligent online counselling framework possesses the potential to modernise social work service delivery within EHWPs, ensuring relevance, accessibility, and sustainability in a technologically advancing environment.

REFERENCES

Acheampong, F.A., Nunoo-Mensah, H. & Chen, W. 2021. Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artificial Intelligence Review*, 54(8):5789-5829.

Agrawal, G., Bertsekas, D. & Liu, H. 2023. Action-based learning for question answering over knowledge graphs. *Information*, 14(6):336-356.

Ahmad, R., Siemon, D., Gnewuch, U. & Robra-Bissantz, S. 2022. Designing personality-adaptive conversational agents for mental health care. *Information Systems Frontiers*, 24(3):923-943.

Akinsola, J.E., Ogunbanwo, A.S., Okesola, O.J., Odun-Ayo, I.J., Ayegbusi, F.D. & Adebiyi, A.A. 2020. *Comparative analysis of software development life cycle models (SDLC)*. Paper delivered at Intelligent Algorithms in Software Engineering: Proceedings of the 9th Computer Science, On-line Conference. https://link.springer.com/chapter/10.1007/978-3-030-51965-0_27 Date of access: 02 Dec. 2023.

Baer, L., Jacobs, D.G., Meszler-Reizes, J., Blais, M., Fava, M., Kessler, R., … Leahy, L. 2000. Development of a brief screening instrument: the HANDS. *Psychotherapy and psychosomatics*, 69(1):35-41.

Baez, M., Daniel, F., Casati, F. & Benatallah, B. 2020. Chatbot integration in a few patterns. *IEEE Internet Computing*, 25(3):52-59.

Bard. 2024. Personal communication. https://bard.google.com/chat/906c82f4e6889f32 Date of access: 22 Jan. 2024.

Bendig, E.B., Schulze-Thuesing, L. & Baumeister, H. 2022. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. *Verhaltenstherapie*, 32(1):64-76.

Betrand, C.U., Onyema, C.J., Benson-Emenike, M.E., Onukwugha, C.G., Onyemauche, U.C. & Kelechi, D.A., Intelligent Learning System Using Interactive Dialogflow and Webhooks. *International Journal of Intelligent Information Systems*, 12(4):54-62.

Bhirud, N., Tataale, S., Randive, S. & Nahar, S. 2019. A literature review on chatbots in the healthcare domain. *International journal of scientific & technology research*, 8(7):225-231.

Boucher, E.M., Harake, N.R., Ward, H.E., Stoeckl, S.E., Vargas, J., Minkel, J., … Zilca, R. 2021. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Review of Medical Devices*, 18(1):37-49.

Brown, J.E. & Halpern, J. 2021. AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare. *SSM-Mental Health*, 1, p100017. https://doi.org/10.1016/j.ssmmh.2021.100017

Brüggen, E. & Willems, P., 2009. A critical comparison of offline focus groups, online focus groups and e-Delphi. *International Journal of Market Research*, 51(3):1-15.

Cahn, J. 2017. *Chatbot: Architecture, Design, & Development*. Philadelphia: University of Pennsylvania (Thesis – PhD).

Christanto, H.J. & Singgalen, Y.A. 2023. Analysis and Design of Student Guidance Information System through Software Development Life Cycle (SDLC) and Waterfall Model. *Journal of Information Systems and Informatics*, 5(1):259-270.

Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P. & D'Alfonso, S. 2023. To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digital Health*, 9, 1-11.

Dankwa-Mullan, I., Scheufele, E.L., Matheny, M.E., Quintana, Y., Chapman, W.W., Jackson, G. & South, B.R. 2021. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. *Journal of Health Care for the Poor and Underserved*, 32(2):300-317.

De Jong, P. 2015. Solution-focused Therapy. In: Corcoran, K. & Roberts, A.R., eds. *Social Workers' Desk Reference*. 3rd ed. New York: Oxford University Press. pp. 268-275.

Department of Communications and Digital Technologies (South Africa). 2024. *South Africa National Artificial Intelligence Policy Framework.* https://www.dcdt.gov.za Date of access: 25 Nov. 2024.

Deshpande, S. & Warren, J. 2021. Self-Harm Detection for Mental Health Chatbots. *Studies in Health Technology and Informatics,* 281:48-52.

Ewing, J.A. 1984. Detecting AlcoholismThe CAGE Questionnaire. *JAMA*, 252(14):1905-1907.

Davidson, J.R., Miner, C.M., De Veaugh-Geiss, J., Tupler, L.A., Colket, J.T. & Potts, N.L.S. 1997. The brief social phobia scale: a psychometric evaluation. *Psychological medicine*, 27(1):161-166.

European Union. 2019. *Ethics guidelines for trustworthy AI*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai Date of access: 27 Jan. 2024.

Finlayson, B.T., Jones, E. and Pickens, J.C. 2023. Solution-focused brief therapy telemental health suicide intervention. *Contemporary family therapy*, 45(1):49-60.

Floridi, L. & Cowls, J. 2019. A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1), art, #1162 doi.org/10.1162/99608f92.8cd550d1

Garcia, E. 2024. Deep Reinforcement Learning for Adaptive Drug Dosage Optimization: Utilizes deep reinforcement learning to optimize drug dosage based on patient response. *Journal of AI-Assisted Scientific Discovery*, 4(1):64-77.

Global Legal Insights. 2023. *AI, Machine Learning & Big Data Laws and Regulations 2023 | South Africa.* https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/south-africa Date of access: 27 Jan. 2024.

Goldberg, S.B., Flemotomos, N., Martinez, V.R., Tanana, M.J., Kuo, P.B., Pace, B.T., … Atkins, D.C. 2020. Machine learning and natural language processing in psychotherapy research: alliance as example use case. *Journal of Counseling Psychology*, 67(4):438–448.

Hamilton, M. 1960. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry,* 23(1):56–62.

Heimler, E. 1990. *Heimler scale of social functioning Rev. IV*. London: Eugene Heimler Trust.

Hudson, W.W. 1982. The clinical measurement package: A field manual. Chicago: Dorsey.

Johnson, E.T., Huck, G., Brooks, J. Moser, E. Blake, J. & Chan, F. 2015. Solution-Focused Brief Therapy. In: Chan, F., Berven, N.L. & Thomas, K.R., eds. *Counseling Theories and Techniques for Rehabilitation and Mental Health Professionals*. New York: Springer. pp. 49-69.

Kasi, V., Keil, M., Mathiassen, L., & Pedersen, K. 2008. The post mortem paradox: A Delphi study of IT specialist perceptions. *European Journal of Information Systems*, 17:62-78.

Kejriwal, M. 2022. Knowledge graphs: a practical review of the research landscape. *Information,* 13:161-178.

Khan, I.A. and Kumari, D., 2021. The Role of Analysis Phase of SDLC for Small Scale Business Application-A Review. *International Journal Of Humanities, Engineering, Science And Management*, 2(01):63-75.

Knott, A. 2023. Job sharing between human professionals and chatbots: How should 'handovers' happen?. *Journal of AI, Robotics & Workplace Automation*, 2(2):121-130.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I. & NeurOx Young People's Advisory Group. 2019. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights*, 11, p.1178222619829083. https://doi.org/10.1177/1178222619829083

Koulouri, T., Macredie, R.D. & Olakitan, D. 2022. Chatbots to support young adults' mental health: an exploratory study of acceptability. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 12(2):1-39.

Landowska, A. 2018. Towards New Mappings between Emotion Representation Models. *Applied Sciences*, 8(2), art. 274. https://doi.org/10.3390/app8020274

Martinengo, L., Lum, E. & Car, J. 2022. Evaluation of chatbot-delivered interventions for self-management of depression: content analysis. *Journal of affective disorders*, 319:598-607.

Mawani, A. & Nderu, L., 2020. Towards an online empathy assisted counselling web application. *EAI Endorsed Transactions on Context-aware Systems and Applications*, 7(22), e4. https://doi-org.10.4108/eai.23-12-2020.167792

Macnab-Holding, C. 2024. *Exploring employees' subjective experience of psychological empowerment after adopting Artificial Intelligence technology*. Stellenbosch: Stellenbosch University. (Dissertation – Masters).

Mead, D. & Mosely, L. 2001. The use of the Delphi as a research approach. *Nurse researcher*, 8(4):4-23.

Meyer, T.J., Miller, M.L., Metzger, R.L. & Borkovec, T.D. 1990. Development and validation of the Penn state worry questionnaire. *Behaviour research and therapy*, 28(6):487-495.

Nosrati, S., Sabzali, M., Heidari, A., Sarfi, T. & Sabbar, S. 2020. Chatbots, counselling, and discontents of the digital life. *Journal of Cyberspace Studies*, 4(2):153-172.

Noyce, R. & Simpson, J. 2016. The Experience of Forming a Therapeutic Relationship from the Client's Perspective: A Metasynthesis. *Psychotherapy Research*. 28(2):281-296.

Oh, K.J., Lee, D., Ko, B. & Choi, H.J. 2017. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. Paper delivered at the 18th IEEE international conference on mobile data management (MDM), Daejeon. https://ieeexplore.ieee.org/document/7962482 Date of access: 02 Dec. 2023.

Pinciroli, F., Justo, J.L.B. & Forradellas, R. 2022. Systematic mapping study: On the coverage of aspect-oriented methodologies for the early phases of the software development life cycle. *Journal of King Saud University-Computer and Information Sciences*, 34(6):2883-2896.

Pothukuchi, A.S., Kota, L.V. and Mallikarjunaradhya, V. 2023. Impact of Generative AI on the Software Development Lifecycle (SDLC). *International Journal of Creative Research Thoughts*, 11(8):227-291.

Rathnayaka, P., Mills, N., Burnett, D., De Silva, D., Alahakoon, D. & Gray, R. 2022. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors*, 22(10):3653-3668.

Roestenburg, W.J.H. 2021. *Analysing technology acceptance by social workers in South Africa*. Potchefstroom: North-West University. (Dissertation – M.B.A.).

Royce, W.W. 1970. Managing the Development of Large Software Systems. *Conference proceedings.* IEEE WESCON. Los Angeles, United States: Western Electronic Show and Convention. pp. 328-388.

Sakurai, E., Kurashige, K., Tsuruta, S., Sakurai, Y., Knauf, R., Damiani, E., … Frati, F. 2020. Embodiment matters: toward culture-specific robotized counselling. *Journal of Reliable Intelligent Environments*, 6:129-139.

Saravanos, A. & Curinga, M.X. 2023. Simulating the Software Development Lifecycle: The Waterfall Model. *Applied System Innovation*, 6(6):108-127.

Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R. & Grant, M. 1993. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, 88(6):791-804.

Schnurr, P., Vielhauer, M., Weathers, F. & Findler, M. 1999. Brief trauma questionnaire. *APA PsycTests*. https://doi.org/10.1037/t07488-000

Selzer, M.L., Vinokur, A. & van Rooijen, L. 1975. A self-administered short Michigan alcoholism screening test (SMAST). *Journal of studies on alcohol*, 36(1):117-126.

Sepahpour, T. 2020. *Ethical considerations of chatbot use for mental health support*. Baltimore: Johns Hopkins University. (Thesis – Masters).

Setyawan, M.Y.H., Awangga, R.M. & Efendi, S.R. 2018. Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot. *Conference proceedings*. 2018 International Conference on Applied Engineering (ICAE, 2018), Batam, Indonesia. Institute of Electrical and Electronics Engineers. pp. 1-5.

Sharma, M.K., 2017. A study of SDLC to develop well engineered software. *International Journal of Advanced Research in Computer Science*, 8(3):520-523.

Shinozaki, T., Yamamoto, Y. & Tsuruta, S., 2015. Context-based counselor agent for software development ecosystem. *Computing*, 97:3-28.

Singh, A., Ramasubramanian, K. & Shivam, S. 2019. Building an enterprise chatbot: work with protected enterprise data using open source frameworks. New York, NY: Apress.

Spitzer, R.L., Kroenke, K., Williams, J.B. & Löwe, B. 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166(10):1092-1097.

Steidl, M., Felderer, M. Ramler, R. 2023. The pipeline for the continuous development of artificial intelligence models—Current state of research and practice. *Journal of Systems and Software*, 199:111615.

Trappey, A.J., Lin, A.P., Hsu, K.Y., Trappey, C.V. & Tu, K.L., 2022. Development of an empathy-centric counseling chatbot system capable of sentimental dialogue analysis. *Processes*, 10(5), 930. https://doi.org/10.3390/pr10050930

Vilaza, G.N. and McCashin, D. 2021. Is the automation of digital mental health ethical? Applying an ethical framework to chatbots for cognitive behaviour therapy. *Frontiers in Digital Health*, 3, p689736. https://doi.org/10.3389/fdgth.2021.689736

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., … Zhang, W., 2022. A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion*, 83:19-52.

Wang, Z. 2023. *Contextual Embedding Representations for Retrieval-Based and Generation-Based Dialogue Systems*. Atlanta: Emory University. (Thesis - PhD).

Watson, D. & Friend, R. 1969. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology,* 33(4):448–457.

Watson, D. & Friend, R. 1969. Social avoidance and distress scale. *Journal of Consulting and Clinical Psychology*, 33:448-457.

Wulf, A. 2020. How to use text messages as a dialogue tool in Solution Focused conversations. In: Dierolf, K., Hogan, D. & van der Hoorn, S., eds. *Solution Focused Practice Around the World*. New York: Routledge. pp. 94-107.

Yildirim, S. & Asgari-Chenaghlu, M. 2021. Mastering Transformers: Build state-of-the-art models from scratch with advanced natural language processing techniques. Packt: Birmingham.

CHAPTER 5: FINDINGS SUMMARY, LIMITATIONS OF THE STUDY, CONCLUSION AND RECOMMENDATIONS

5.1. SUMMARY

Chapter 5 provides a synthesis of the study's findings and presents a comprehensive summary, conclusion, and recommendations for implementing an AI-driven counselling framework within social services, specifically the EHWP. The following sub-sections provide a chapter-by-chapter account of the research conducted, underscoring the foundational insights and empirical evidence supporting the framework.

5.1.1. Chapter 1: Introduction to the research

This chapter outlined the context and rationale for developing an intelligent online counselling framework for an EHWP. The chapter detailed the significance of integrating AI in social work to modernise the field and enhance service delivery. The problem statement identified the deficiency in the use of AI, particularly in the field of social work and EHWP.

Key insights emphasised the evolution of AI therapy agents, the increasing use of HITs in the modernisation of healthcare and the importance of social work integrating these technologies to enhance service delivery and clinical outcomes.

The chapter laid the foundation for the subsequent chapters by posing the research question *What aspects must be considered in developing an AI generative chatbot counselling framework for a social work EHWP (Lyra SA)?* The research sub-questions that aimed to inform the research question include:

- What literature regarding the use of AI in psychosocial health is available?;
- How do live counselling clinicians currently address service user (client) problems through text-based live chat services?;
- What are the textual trends within current live chat services data?;
- What components of the conceptual AI generative chatbot counselling framework inform a multiple-response AI algorithm?; and
- Which inputs from subject matter experts could improve the proposed AI generative chatbot counselling framework?

To answer the research questions, a unique approach to social work research was devised. The study drew upon a multi-phased mixed-method research approach, which integrated both qualitative and quantitative methods to uncover comprehensive insights into AI counselling frameworks.

The sequence of the methodological approach was to conduct an initial qualitative analysis, followed by a quantitative analysis. This approach ensured that insights from qualitative data could inform the subsequent quantitative analysis.

The research phases consisted of a literature study to conduct an extensive review of existing frameworks, technologies and services and to establish a theoretical discourse. This review identified key concepts and current research on the use of AI in therapeutic services, EHWPs and social work.

The qualitative phase of the study (Phase 1) employed thematic analysis of a sample of live chat transcripts. It identified interactive pathways and therapeutic approaches implemented by human counsellors during live chat counselling sessions.

The quantitative phase of the study (Phase 2) employed frequency analysis of the live chat dataset by drawing upon codes identified in the qualitative phase. Lexicon-based sentiment analysis was conducted to quantify data by identifying sentiment and word frequencies.

Data integration (Phase 3) of the study converged and interpreted the qualitative and quantitative findings to inform comprehensive conclusions. The converged data formed the basis for the development of the conceptual framework.

The development of the conceptual framework (Phase 4) was conducted by drawing upon the principles of the SDLC, informed by the Multinomial Naïve Bayes Theorem. The conceptual framework was refined through the Delphi study methodology, which involved feedback and the integration of feedback from social workers, registered counsellors, psychologists and ICT experts.

Throughout the study, ethical considerations remained prioritised and proved to pose minimal risks to participants. The researcher also maintained confidentiality by anonymising data for analysis.

5.1.2. Chapter 2: Literature Study

The chapter provided a comprehensive literature study of existing literature regarding the use of AI in mental healthcare, chatbots and their application in social work. The literature study underscored the benefits and challenges of implementing AI-driven interventions in mental healthcare, thereby providing the premise for empirical research in the subsequent chapters.

The chapter explored the evolution of AI from Turing's code-breaking technologies during WWII to the development of sophisticated AI applications, such as ChatGPT and others. The focus predominantly remained on user engagement in natural language processing (NLP)-driven conversations, which are increasingly being employed across various industries.

The relevance of AI to social work was explored through the context of the 4IR, which is empowered by the IoT, transforming traditional service delivery. These technological advancements necessitate social work to modernise, in order to remain relevant in a rapidly advancing technological landscape. AI counselling chatbots are but one application that can support social work and EHWPs by providing technologically advanced and accessible mental healthcare services.

The benefits of AI counselling chatbots were explored. The researcher found that these applications are available 24 hours a day, seven days a week at a reduced cost with increased efficiency in service delivery. These applications have the potential to provide personalised interactions and support while reducing the stigma and other hindrances associated with seeking and receiving mental healthcare services.

AI counselling chatbots have been found to pose limitations in terms of technical issues and data security concerns. The potential for biased responses, limited understanding of complex human emotions and difficulties in handling nuanced conversations while providing context-specific responses were explored in-depth.

The holistic features of AI counselling chatbots have been found to necessitate the mimicking of human conversational patterns and behaviours, whereby the chatbot can understand natural language and provide empathic, personalised responses. The incorporation of sentiment analysis and emotional understanding are critical elements of this feature.

Various methodologies in AI chatbot development were identified as being cardinal to success. These methodologies include text analysis, data pre-processing, input embedding, NLP, Naïve Bayes classification, ML, sentiment analysis and ANNs.

The role of big data in the development of AI counselling chatbots was investigated. Findings indicated that big data is essential in the training of AI counselling chatbots to improve accuracy. In-depth analysis of big data enables the tailoring of chatbot responses and the improvement of service delivery. However, specific language registers are required to improve AI counselling chatbots' understanding and generation of human-like responses.

The role of AI chatbots in social work and EHWPs was explored. Findings indicated that AI chatbots possess the potential to supplement social work services by providing accessible mental healthcare services. AI chatbots have proven to enhance the efficiency of EHWPs, especially amongst younger technologically inclined users.

A suitable clinical theoretical framework was investigated. The SFBT model was identified as a viable option for an AI-driven counselling framework.

Focus was placed on the ethical considerations in the deployment of an AI-driven framework. It was explicitly found that an AI-driven counselling framework must adhere to stringent ethical principles such as beneficence, non-maleficence, autonomy, justice and explicability. Biases must be avoided while maintaining professional standards. The protection of data from a systems perspective is of the utmost importance.

The literature study concluded that social work practitioners must employ digital tools when striving to modernise their approaches and remain relevant in the 4IR. Collaboration between IT professionals and mental healthcare practitioners is crucial in achieving this ideal.

Continued cross-disciplinary and multi-disciplinary research is required for the integration of AI into social work practice. Digital competencies should be incorporated into social work curricula to prepare future social workers for the demands of an ever-increasingly digitalised world.

5.1.3. Chapter 3: Empirical Findings

A multi-phased mixed-method approach was implemented to analyse live chat counselling transcripts provided by Lyra SA. The qualitative thematic analysis and quantitative textual analysis uncovered pertinent trends and patterns in the utilisation of live chat counselling services, user concerns and the implementation of current text-based therapeutic approaches.

The thematic analysis of 90 live chat transcripts highlighted numerous pertinent themes and sub-themes. These included the flow/process flow of live chat sessions, user problems/challenges, counsellor approaches, linguistic trends and technical aspects.

The flow/process flow of live chat sessions found a similar process flow across all analysed transcripts. Specifically, that the therapeutic session commenced with the exchange of greetings, that user biographical information was gathered, problems were explored, goal setting was discussed and the therapeutic intervention was shared.

Users connected and sought services relating to a variety of problems/challenges. These included, inter alia, personal development, dysfunction, psychological challenges, emotional wellbeing, career/work-related issues, interpersonal relationships, family dynamics, requests for help/support, suicidal ideation and other diverse challenges.

Live chat counsellors employed three main approaches to the sessions, which included advice and support whereby actionable advice, encouragement and suggestions for coping mechanisms were provided. Live chat counsellors primarily utilised the psychological paradigms of CBT and SFBT. Counsellors applied various therapeutic techniques such as empathy, affirmation, goal setting and challenging negative thought patterns (to note a few) in their therapeutic approach.

The analysis of linguistic trends identified the use of informal communication, emotional interaction and expressive interaction. It also discovered trends regarding written expressions, various degrees of tones of politeness, formal language and specialised language.

Various technical issues that periodically interrupted the therapeutic sessions were identified. These included counsellors' shifts ending, system errors, loadshedding, network issues and users' disconnections.

Sentiment analysis was conducted whereby 47% of sentiments were classified as negative, 25% were classified as neutral and 28% were classified as positive. The sentiment analysis underscored the complexity of live chat sessions and illustrated the importance of addressing both positive and negative emotions in delivering support services.

5.1.4. Chapter 4: Proposed Product: An Intelligent Online Counselling Framework for an Employee Health and Wellness Programme (EHWP)

The argument for the need for social work to incorporate digital technologies, adopting data management tools, digital interventions and digital competencies was made in the preceding chapters. In this chapter, the empirical findings provided the basis for the proposed online counselling framework.

The framework was developed to support social work practice within an EHWP by providing a supplementary AI-driven counselling framework. The chapter detailed the architectural design, data training, algorithms, ethical considerations and operational aspects of such a framework. The framework was subject to the scrutiny of a Delphi panel. The Delphi panel's recommendations were analysed and incorporated into the proposed product.

The proposed framework's goals and objectives were based on empirical evidence. Goals included accessible counselling, addressing user-specific concerns, conducting evidence-based assessments and the implementation of sound psychological frameworks.

The objectives of the AI framework were formulated as a result of the analysis phase. This phase aimed to identify problems and needs and propose solutions by implementing clinical and technical methodologies to ensure optimal clinical outcomes for service users.

The clinical technical requirements for the proposed product were identified as essential elements to ensure its functionality and effectiveness. These requirements encompass both clinical and technical aspects necessary for the development of the intelligent online counselling framework.

By integrating these rigorous prerequisites, the framework aims to meet the specific needs of users within the EHWP. These requirements include NLP, emotion recognition and affective computing, contextual, awareness, knowledge base integration, dynamic risk assessment and crisis response, adaptability and learning. They further include therapeutic approaches, goal

setting, progress tracking, referral mechanisms, cultural sensitivity, personalisation, empathy and tone, user engagement strategies and community resources or support groups.

The proposed product has, at its core, the aim of enhancing social work practice within EHWPs, by integrating state-of-the-art digital technologies into the practice of social work. This chapter synthesised the findings of theoretical- and empirical research findings to recommend a framework designed to supplement and enhance the accessibility, effectiveness and personalisation in the modern workplace.

5.2. LIMITATIONS OF THE STUDY

5.2.1. Subjectivity in thematic analysis

As a qualitative research method, thematic analysis inherently relies on the researcher's interpretation in the identification and synthesis of patterns within data. Rigorous procedures, such as systematic coding, cross-referencing of themes, and the use of software to assist in identifying recurrent patterns were applied to enhance the reliability of this study. It is acknowledged that, despite these measures, qualitative analyses are influenced by the perspective and experiences of the researcher, which may subtly shape the thematic categorisations and interpretations.

While every effort was made to mitigate subjective bias—through the use of detailed protocols, methodological triangulation, and iterative consultations—the possibility of residual subjectivity remains a characteristic of this approach. Such nuances, however, are not indicative of methodological shortcomings but reflect the depth and complexity inherent in qualitative research. This study views these interpretative layers as enriching the findings, providing a nuanced understanding of the data that aligns with the broader objectives of exploring innovative frameworks in social work and AI integration.

5.2.2. Adapting the research focus

The initial aim of this study was to design and develop an AI counselling application tailored for EHWPs. However, guidance from the HREC necessitated a shift in focus. Ethical considerations, particularly concerning participant safety and data privacy, led to the recommendation that the study focus on developing a conceptual framework rather than an operational application.

While this adjustment ensured ethical compliance, it presented a challenge to the broader goal of advancing AI integration into social work. The framework developed is comprehensive and provides a valuable theoretical foundation for future innovations in the field. However, the absence of a developed application reduces the opportunity for direct application at this stage. Unlike psychology, where AI tools such as chatbots are actively implemented and tested, social work lacks similar advancements that could demonstrate practical integration into existing practices.

This limitation underscores the need for ongoing investment and multidisciplinary collaboration to move from conceptual frameworks to real-world applications. Such efforts will be critical to ensuring that the potential benefits of AI can be realised within social work and EHWPs, allowing the profession to remain relevant and innovative in the rapidly evolving landscape of mental health services.

5.2.3. Limited literature on AI in social work

At the commencement of this study, a critical challenge was identified, namely the limited availability of scholarly literature addressing the application of AI within the domain of social work. While the use of AI in disciplines such as psychology, healthcare, and education has been extensively documented, social work lags in this regard. Early literature searches revealed a wealth of research on AI's integration into psychological services, including the use of chatbots and other AI-driven therapeutic tools. However, comparable studies or applications tailored to social work were scarce.

This scarcity of resources presented a notable limitation throughout the research process. The absence of established frameworks, case studies, or empirical evidence specific to social work necessitated reliance on adapting insights from other disciplines. While this approach provided valuable guidance, it also highlighted the unique needs and nuances of social work that require dedicated research.

The limited literature also underscored the need for this study. By addressing the gap, the research contributes to an emerging area of inquiry, laying foundational groundwork for the integration of AI into social work practice. However, it is evident that more focused investigations, resources, and collaborations are needed to foster a comprehensive understanding of AI's potential within the profession. This gap not only constrained the study but also reaffirmed its importance in positioning social work as a field that can embrace and benefit from technological advancements.

5.3. CONCLUSION

This thesis investigated the development of an intelligent online counselling framework for an EHWP. Emphasis was placed on the potential influence of AI in the field of social work. The advancements of the 4IR and the need to modernise social work in an increasingly advancing digital world served as motivation for the study.

This study integrated a multi-phased mixed-method approach to investigate text-based live chat services. The thematic and textual analysis of live chat transcripts provided insight into the nature of client concerns and the counselling approaches employed by mental healthcare practitioners. The insights gained from the analyses were critical in the design of a conceptual AI framework,

which aims to supplement human counsellors to enhance the accessibility and efficiency of mental healthcare services.

The proposed framework draws upon AI technologies to theorise a generative counselling chatbot that is capable of delivering immediate, consistent and empathic support to users.

Various limitations of traditional live chat services, such as differing engagement strategies/styles, counsellor fatigue, shift changes and other elements are addressed by this AI-driven framework. The framework possesses the potential to improve service delivery within the EHWP by providing a scalable and reliable support system.

The practical relevance and potential effectiveness were validated through the feedback from a Delphi panel of experts. The integration of AI into social work practice has the potential to represent a significant advancement in its capacity to meet the evolving needs of service users.

The role of AI in modernising social work, particularly in the field of mental healthcare, has been demonstrated within this thesis. A viable solution for enhancing the effectiveness of mental healthcare is demonstrated through the proposed intelligent online counselling framework. Timely consistent and scalable care through AI-driven support can supplement human counsellors in ultimately improving the wellbeing of service users.

The successful implementation of the proposed framework would require interdisciplinary and multi-disciplinary collaboration, continuous improvement and the meticulous consideration of ethical guidelines. This study provides the opportunity for future research to focus on piloting the framework in a real-world setting, whereby the broader application of AI in social work can be explored.

By embracing the advancements of the 4IR, social work has the opportunity to enhance its relevance and effectiveness. In doing so, practitioners can ensure that social work continues to play a vital role in supporting mental health and wellbeing in an advancing digital world.

5.4. RECOMMENDATIONS

The following sub-sections present recommendations for integrating the proposed AI counselling framework into social work practices within an EHWP.

5.4.1. Practical applications

- Implementation of the AI counselling frameworks in EHWP: It is recommended that the proposed framework be considered for development and implementation into the service offerings of Lyra SA. This could be introduced as a supplementary tool to existing services.
- Training and development: It is recommended that social workers and social work educators receive thorough training on the integration and application of AI in traditional

social work practice. Research should be conducted to develop guidelines and best practice standards for the integration of AI into traditional social work practice.

- User experience optimisation: Continuous feedback from service users should be gathered in a bid to enhance the framework's interaction quality, responsiveness and quality of service. Updates and refinements based on empirical research should regularly be implemented.

5.4.2. Policy recommendations

- Ethical considerations and data privacy: Clear policies regarding data privacy and confidentiality for the proposed framework should be developed. Furthermore, ethical considerations and guidelines for the use of AI in social work should be developed, whereby compliance with legislative and regulatory requirements is ensured.
- Multi-disciplinary collaboration: Collaboration between social workers, mental healthcare professionals, AI developers, application developers and policymakers should be established to ensure that AI frameworks meet the needs and requirements of all stakeholders. Multidisciplinary research should be encouraged to refine and expand AI capabilities in social work.

5.4.3. Future research opportunities

- Effectiveness of AI in various contexts: Longitudinal studies need to be conducted to assess the impact of AI in various social work settings.
- Advancements in AI technology: Advancements in AI technology, such as ML and NLP need to be investigated to determine its potential benefits regarding its integration into social work. The potential of multimodal AI (combination of text, voice and visual interactions) as an attempt to provide a more holistic therapeutic experience should be investigated.
- Impact on social work practice: The impact of AI integration on the workload and job satisfaction of social workers should be investigated. The potential of AI to mitigate burnout and enhance the overall efficiency of social workers warrants thorough investigation.
- AI algorithm development: AI algorithms tailored for the use in social work and EHWPs should be prioritised for future research and development. The creation of these algorithms should align with the ethical, socio-cultural and contextual principles unique to social work.
- Collaborative product development: Multidisciplinary collaboration between social workers, mental healthcare practitioners and policy makers should be fostered to develop intelligent counselling applications and platforms for use in social work and EHWPs.

Through the consideration of the noted recommendations, the science and field of social work could potentially leverage AI technology to enhance the provision of services, improve service

outcomes and remain relevant in a rapidly evolving technological landscape. The proposed framework can be considered as a foundation for future research and development, contributing to innovative solutions in social work service delivery.

SECTION B

COMPLETE LIST OF REFERENCES

Abbott, S., Ade-Ibijola, A. & Roestenburg.W.J.H. 2021. Big data in social science research. In: Fouche', C.B., Strydom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 181-198.

Abd-alrazaq, A.A., Alajlani, M., Alalwan, A.A., Bewick, B.M., Gardner, P. & Househ, M. 2019. An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, 132, art. #103978. doi:10.1016/j.ijmedinf.2019.103978

Abdul-Kader, S.A. & Woods, J. 2015. Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications*, 6(7):72-80.

Acheampong, F.A., Nunoo-Mensah, H. & Chen, W. 2021. Transformer models for text-based emotion detection: a review of BERT-based approaches. Artificial Intelligence Review, 54(8):5789-5829.

Adamopoulou, E. & Moussiades, L. 2020. *An overview of chatbot technology*. Paper delivered at IFIP International Conference on Artificial Intelligence Applications and Innovations, Cham. https://link.springer.com/chapter/10.1007/978-3-030-49186-4_31 Date of access: 18 Dec. 2022.

Adams, D. 1979. *The hitchhiker's guide to the galaxy*. New York, NY: Harmony Books.

Adams, D. 1980. *The Restaurant at the End of the Universe*. New York, NY: Harmony Books.

Adams, D. 1982. *Life, the Universe and Everything*. New York, NY: Harmony Books.

AFP Wire Service. 2023. LOOK: ChatGPT bot passes US law school exam. *IOL,* 28 Jan. https://www.iol.co.za/technology/software-and-internet/look-chatgpt-bot-passes-us-law-school-exam-7eade96a-fde5-401d-9536-04e18d2fd0a1 Date of Access: 28 Jan. 2023.

Agrawal, G., Bertsekas, D. & Liu, H. 2023. Action-based learning for question answering over knowledge graphs. *Information*, 14(6):336-356.

Ahmad, N.A., Che, M.H., Zainal, A., Abd Rauf, M.F. & Adnan, Z. 2018. Review of chatbots design techniques. *International Journal of Computer Applications,* 181(8):7-10.

Ahmad, R., Siemon, D., Gnewuch, U. & Robra-Bissantz, S. 2022. Designing personality-adaptive conversational agents for mental health care. *Information Systems Frontiers*, 24(3):923-943.

Akinsola, J.E., Ogunbanwo, A.S., Okesola, O.J., Odun-Ayo, I.J., Ayegbusi, F.D. & Adebiyi, A.A. 2020. *Comparative analysis of software development life cycle models (SDLC)*. Paper delivered at Intelligent Algorithms in Software Engineering: Proceedings of the 9th Computer Science , On-

line Conference. https://link.springer.com/chapter/10.1007/978-3-030-51965-0_27 Date of access: 02 Dec. 2023.

Andersson G. 2016. Internet-delivered psychological treatments. *Annu Rev Clin Psychol*, 12:157–79.

Anney, V.N. 2014. Ensuring the quality of the findings of qualitative research: looking at trustworthiness criteria. *Journal of emerging trends in educational research and policy studies*, 5(2):272-281.

Anthony, K. 2003. The use and role of technology in counselling and psychotherapy. In: Goss, S. & Anthony, K., eds. *Technology in counselling and psychotherapy: A practitioner's guide*. London: Palgrave. pp. 13-34.

Archard, P.J. 2021. The psychoanalytically-informed interview in social work research. *Journal of Social Work Practice*, 35(2):191-203.

Argyropoulou, K. & Kaliris, A. 2018. From career decision-making to career decision-management: New trends and prospects for career counseling. *Advances in Social Sciences Research Journal*, 5(10):483-502.

Ashton, K. 2009. That 'internet of things' thing. *RFID journal*, 22(7):97-114.

Augar, N. & Zeleznikow, J. 2014. Developing online support and counseling to enhance family dispute resolution in Australia. *Group Decision and Negotiation*, 23:515-532.

Ausman, M.C. 2019. Artificial Intelligence's Impact on Mental Health Treatments. Paper delivered at the AAAI/ACM Conference on A.I., Ethics, and Society (AIES' 19), Honolulu. https://doi.org/10.1145/3306618.3314311 Date of access: 10 Mar. 2022.

Aveyard, H. 2010. *Doing a literature review in health and social care: A practical guide.* 2nd ed. Berkshire, Great Britain: Open University Press.

Babbie, E.R. 2017. The Basics of Social Research. 7th ed. Boston: Cengage Learning.

Baer, L., Jacobs, D.G., Meszler-Reizes, J., Blais, M., Fava, M., Kessler, R., … Leahy, L. 2000. Development of a brief screening instrument: the HANDS. *Psychotherapy and psychosomatics*, 69(1):35-41.

Baez, M., Daniel, F., Casati, F. & Benatallah, B. 2020. Chatbot integration in few patterns. *IEEE Internet Computing*, 25(3):52-59.

Balcombe, L. & De Leo, D. 2022. Evaluation of the Use of Digital Platforms and Digital Mental Health Interventions: Scoping Review. *International Journal of Environmental Research and Public Health*. 20(1):362.

Baldwin, R.E. 2019. *The Globotics Upheaval: Globalization, Robotics, and the Future of Work*. New York, NY: Oxford University Press.

Ball, R. & Rague, B. 2022. *The Beginner's Guide to Data Science*. Cham: Springer International Publishing.

Bansal, H. & Khan, R. 2018. A review paper on human computer interaction. *International Journal of Advanced Research in Computer Science and Software Engineering,* 53(8), art. #8 doi:10.23956/ijarcsse.v8i4.630

Bard. 2024. Personal communication. https://bard.google.com/chat/906c82f4e6889f32 Date of access: 22 Jan. 2024.

Barker, S. 2021. Deep Learning Chatbot: Everything You Need to Know. https://shanebarker.com/blog/deep-learning-chatbot/ Date of access: 30 Apr. 2022.

Barnes, J., Conrad, K., Demont-Heinrich, C., Graziano, M., Kowalski, D., Neufeld, J., Zamora, J. & Palmquist, M. 2012. *Generalizability and transferability*. http://writing.colostate.edu/guides/guide.cfm?guideid=65 Date of access: 11 Mar. 2022.

Basha, S.M. & Fathima, A.S. 2023. *Natural Language Processing: Practical Approach*. Karnataka: MileStone Research Publications.

Bassil, Y. 2012. A Simulation Model for the Waterfall Software Development Life Cycle. *International Journal of Engineering & Technology (iJET)*, 2(5):2049-3444.

Bauer, J. 2021. *The Transformative Self: Personal Growth, Narrative Identity, and the Good Life*. New York, NY: Oxford University Press.

Beg, M.J., Verma, M., M, V.C.K.M. & Verma, M.K. 2024. Artificial Intelligence for Psychotherapy: A review of the current state and future directions. *Indian Journal of Psychological Medicine.* (July, 28).

Bendig, E.B., Schulze-Thuesing, L. & Baumeister, H. 2022. The next generation: chatbots in clinical psychology and psychotherapy to foster mental health–a scoping review. *Verhaltenstherapie*, 32(1):64-76.

Benyera, E. 2021. The Fourth Industrial Revolution and the Recolonisation of Africa: The Coloniality of Data. New York: Routledge.

Bergman, S. & Humble, N. 2022. Sentiment analysis for emotional navigation in written communication: What support do autistic people need? *Artificial Intelligence and Robotics*, 4(1):6-11.

Bernard, H. R. 1996. *Qualitative data, quantitative analysis*. The Cultural Anthropology Methods Journal, 8(1):9–11.

Betrand, C.U., Onyema, C.J., Benson-Emenike, M.E., Onukwugha, C.G., Onyemauche, U.C. & Kelechi, D.A., Intelligent Learning System Using Interactive Dialogflow and Webhooks. International Journal of Intelligent Information Systems, 12(4):54-62.

Bhadane, C., Dalal, H. & Doshi, H., 2015. Sentiment analysis: Measuring opinions. *Procedia Computer Science*, 45:808-814.

Bhangdia, C., Jadhav, S., Gadgil, T., Kumari, A. & Dasari, M. 2022. Sentiment Analysis using Chatbot and Mental Health Tracker. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1):131-136.

Bhirud, N., Tataale, S., Randive, S. & Nahar, S. 2019. A literature review on chatbots in healthcare domain. *International journal of scientific & technology research*, 8(7):225-231.

Binder, P., Moltu, C., Hummelsund, D., Sagen, S.H. & Holgersen, H. 2011. Meeting an adult ally on the way out into the world: Adolescent patients' experiences of useful psychotherapeutic ways of working at an age when independence really matters. *Psychotherapy Research*, 21(5):554–566.

Blei, D.M., Ng, A.Y. & Jordan, M.I. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993-1022.

Bonaccorso, G. 2017. *Machine learning algorithms*. Birmingham: Packt Publishing.

Boslaugh, S. 2007. *Secondary Data Sources for Public Health: A Practical Guide*. St Louis: Washington University.

Botma, Y., Greeff, M., Mulaudzi, F.M. & Wright, S.C.D. 2010. *Research in health science*. Cape Town: Heinemann, Pearson Education South Africa.

Boucher, E.M., Harake, N.R., Ward, H.E., Stoeckl, S.E., Vargas, J., Minkel, J., Parks. A.C. & Zilca, R. 2021. Artificially intelligent chatbots in digital mental health interventions: a review. *Expert Review of Medical Devices*, 18:37-49.

Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology,* 3:77-101.

Bridgman, L.A. & Lazar, B.S. 2019. A Systematic Literature Review of the use of Artificial Intelligence (AI) or Mobile Applications for Depression Therapy. *International Journal of Academic Health and Medical Research*, 3(7):18-24.

Brooke, S. 2017. *3 Revolutionary Ways of Using Chatbots for Big Data*. https://chatbotslife.com/3-revolutionary-ways-of-using-chatbots-for-big-data-c0a9076171c3 Date of access: 10 Jan. 2023.

Brown, J.E. & Halpern, J. 2021. AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare. *SSM-Mental Health*, 1, p100017. https://doi.org/10.1016/j.ssmmh.2021.100017

Brüggen, E. & Willems, P., 2009. A critical comparison of offline focus groups, online focus groups and e-Delphi. *International Journal of Market Research*, 51(3):1-15.

Cahn, J. 2017. *Chatbot: Architecture, Design, & Development*. Philadelphia: University of Pennsylvania (Thesis – PhD).

Casillo, M., Clarizia, F., D'Aniello, G., De Santo, M., Lombardi, M. & Santaniello, D. 2020. CHAT-Bot: A cultural heritage aware teller-bot for supporting touristic experiences. Pattern *Recognition Letters*, 131:234-243.

Catelli, R, Pelosi, S. & Esposito, M. Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. *Electronics,* 11, art. #374. doi: 10.3390/electronics11030374

Chalawila, I. & Muchanga, M., 2022. Challenges experienced by postgraduate candidates in the application of conceptual frameworks in scientific Research. *International Journal of Scientific Research and Management (IJSRM)*, 10(2):2174-2183.

Chatterjeea, A., Guptaa, U., Chinnakotlab, M.K., Srikanthb, R., Galleyc, M. & Agrawala, P. 2019. Understanding Emotions in Text Using Deep Learning and Big Data. *Computers in Human Behavior*, 93:309-317.

Chaves, A.P. & Gerosa, M.A. 2022. *The Impact of Chatbot Linguistic Register on User Perceptions: A Replication Study*. Paper delivered at the Chatbot Research and Design: 5th International Workshop, Virtual Event. https://link.springer.com/chapter/10.1007/978-3-030-94890-0_9 Date of access: 08 Apr. 2023.

Chen, Y., Chu, H. & Wang, P. 2021. Employee assistance programs: a meta-analysis. *Journal of employment counselling*, 58:144-166.

Chen, Z. & Chen, J. 2022. Decision-Making Model of Social Work Specialization Process Using Big Data Analysis and Chaos Cloud. *Wireless Communications and Mobile Computing*, 1-11.

Cherry, K. 2021. *The Pros and Cons of Online Therapy.* https://www.verywellmind.com/advantages-and-disadvantages-of-online-therapy-2795225#:~:text=Online%20access%20makes%20it%20easier,help%20you%20become%20psychologically%20stronger Date of access: 06 Apr. 2022.

Chivers, T. 2019. *The AI Does Not Hate You: Superintelligence, Rationality and the Race to Save the World*. London: Weidenfeld & Nicolson.

Christanto, H.J. & Singgalen, Y.A. 2023. Analysis and Design of Student Guidance Information System through Software Development Life Cycle (SDLC) and Waterfall Model. *Journal of Information Systems and Informatics*, 5(1):259-270.

Cipoletta, S., Entilli, L., Bettio, F. & De Leo, D. 2021. Live-Chat Support for People Bereaved by Suicide: A Thematic Analysis. *Crisis,* 43(2):98-104.

Coghlan, S., Leins, K., Sheldrick, S., Cheong, M., Gooding, P. & D'Alfonso, S. 2023. To chat or bot to chat: Ethical issues with using chatbots in mental health. *Digital Health*, 9, 1-11.

Colby, K.M., Weber, S. & Hilf, F.D. 1971. Artificial paranoia. *Artificial Intelligence*, 2(1):1-25.

Coria, M.D., Arevalo, J.C.M., Hilario-Rivas, J.L., Hilario-Cárdenas, J.R. & Prado-Juscamaita, J.I. 2021. Supervised Sentiment Analysis Algorithms. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(14):2000-2012.

Cowls, J., Tsamados, A., Taddeo, M. & Floridi, L. 2021. A definition, benchmark and database of AI for social good initiatives. *Nat Mach Intell*, 3:111–115.

Creswell, J.W. & Creswell, J.D. 2018. *Research design: qualitative, quantitative and mixed methods approach.* 5th ed. London: Sage Publications.

Creswell, J.W. 2014. *Research design: qualitative, quantitative and mixed methods approach.* London: Sage Publications.

Crutzen, R. 2011. An Artificially Intelligent Chat Agent that Answers Adolescents' Questions Related to Sex, Drugs and Alcohol: An Exploratory Study. *Journal of Adolescent Health,* 48(5): 514-19.

Daniel, F., Cappiello, C. & Benatallah, B. 2019. Bots acting like humans: Understanding and preventing harm. *IEEE Internet Computing*, 23(2):40-49.

Dankwa-Mullan, I., Scheufele, E.L., Matheny, M.E., Quintana, Y., Chapman, W.W., Jackson, G. & South, B.R. 2021. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. Journal of Health Care for the Poor and Underserved, 32(2):300-317.

Davidson, J.R., Miner, C.M., De Veaugh-Geiss, J., Tupler, L.A., Colket, J.T. & Potts, N.L.S. 1997. The brief social phobia scale: a psychometric evaluation. *Psychological medicine*, 27(1):161-166.

De Jesus, A. 2019. *Chatbots for Mental Health and Therapy – Comparing 5 Current Apps and Use Cases.* https://emerj.com/ai-application-comparisons/Chatbots-mental-health-therapy-comparing-5-current-apps-use-cases/ Date of access: 10 Mar. 2022.

De Jong, P. 2015. Solution-focused Therapy. In: Corcoran, K. & Roberts, A.R., eds. *Social Workers' Desk Reference*. 3rd ed. New York: Oxford University Press. pp. 268-275.

de Meyrick, J. 2003. The Delphi method and health research. *Health Education,* 103(1):7-16.

De Vos, A.S. 2005. Qualitative data analysis and interpretation. In: De Vos, A.S., Strydom, H., Fouché, C.B. & Delport, C.S.L. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 333-354.

Deloitte Development. 2018. *The Fourth Industrial Revolution is here—are you ready?* Deloitte Insights.

Demir, S.B. & Pismek, N. 2018. A Convergent Parallel Mixed-Methods Study of Controversial Issues in Social Studies Classes: A Clash of Ideologies. *Educational sciences: theory & practice*, 18(1):119-149.

Denecke, K., Abd-Alrazaq, A., & Househ, M. 2021. Artificial Intelligence for Chatbots in Mental Health: Opportunities and Challenges. In: Househ, M., Borycki, E., Kushniruk, A. eds. *Multiple Perspectives on Artificial Intelligence in Healthcare*. Cham: Springer. pp. 115-128.

Department of Communications and Digital Technologies (South Africa). 2024. South Africa National Artificial Intelligence Policy Framework. https://www.dcdt.gov.za Date of access: 25 Nov. 2024.

Deshpande, S. & Warren, J. 2021. Self-Harm Detection for Mental Health Chatbots. *Studies in Health Technology and Informatics,* 281:48-52.

Dhyani, M. and Kumar, R., 2021. An intelligent Chatbot using deep learning with Bidirectional RNN and attention model. *Materials today*, 34(3):817-824.

Divayana, D.G.H., Sappaile, B.I., Pujawan, I., Dibia, I.K., Artaningsih, L., Sundayana, I. & Sugiharni, G.A.D. 2017. An evaluation of the instructional process of the expert system course program by using the mobile technology-based CSE-UCLA model*. International Journal of Interactive Mobile Technologies*, 11(6):18-31.

Donohoe, H., Stellefson, M. & Tennant, B. 2012. Advantages and Limitations of the e-Delphi Technique: Implications for Health Education Researchers. *American Journal of Health Education*, 43(1):38-46.

Doyle, L., McCabe, C., Keogh, B., Brady, A. & McCann, M. 2019. An overview of the qualitative descriptive design within nursing research. *Journal of Research in Nursing*. 25(5):443-455.

Doyle, L., Brandy, A. & Byrne, G. 2016. An overview of mixed methods research – revisited. *Journal of Research in Nursing,* 21(8):623-635.

El-Ansari, A. & Beni-Hssane, A. 2023. Sentiment Analysis for Personalized Chatbots in E-Commerce Applications. *Wireless Personal Communications*, 129(3):1623-1644.

European Union. 2019. *Ethics guidelines for trustworthy AI*. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai Date of access: 27 Jan. 2024.

Ewing, J.A. 1984. Detecting AlcoholismThe CAGE Questionnaire. *JAMA*, 252(14):1905-1907.

Fainchtein, L. 2020. 28 June. *Generative vs Retrieval Based Chatbots: A Quick Guide* [Blog post]. https://blog.cloudboost.io/generative-vs-retrieval-based-chatbots-a-quick-guide-8d19edb1d645 Date of access: 01 May. 2022.

Fang, L., Tarshis, S., McInroy, L. & Mishna, F. 2018. Undergraduate Student Experiences with Text-Based Online Counselling. *British Journal of Social Work*, 48(6):1774-1790.

Fareed, S.B.S. 2021. *API Pipeline for Visualising Text Analytics Features of Twitter Texts.* Paper delivered at the 2021 International Conference of Women in Data Science at Taif University (WiDSTaif), Al Hawiyah, Saudi Arabia. https://ieeexplore.ieee.org/document/9430213 Date of access: 30 Apr. 2022.

Fava, N.M., Schuler, B.R. & Kelly, M.S. 2016. The incorporation of big data in social work practice: Current efforts and future directions. *Journal of Social Work Education*, 52(3):316-328.

Fernández, E.N. & Mairal, J.B. 2017. Behavioral activation versus cognitive restructuring to reduce automatic negative thoughts in anxiety generating situations. *Psicothema*, 29(2):172-177.

Fetters, M.D., Curry, L.A. & Creswell, J.W. 2013. Achieving Integration in Mixed Methods Designs—Principles and Practices. *Health Services Research*, 48(6):2134-2156.

Finlayson, B.T., Jones, E. and Pickens, J.C. 2023. Solution focused brief therapy telemental health suicide intervention. *Contemporary family therapy*, 45(1):49-60.

Fiske, A., Henningsen, P. and Buyx, A. 2019. Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of medical Internet research*, 21(5), e13216. https://www.jmir.org/2019/5/e13216/

Floridi, L. & Cowls, J. 2019. A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1), art, #1162 doi.org/10.1162/99608f92.8cd550d1

Folstad, A., Brandtzaeg, P.B., Feltwell, T., Effie, L.C., Tscheligi, M. & Luger, E.A. 2018. *Chatbots for Social Good.* Paper delivered at the C.H.I. Conference on Human Factors in Computing Systems, Montreal. https://sintef.brage.unit.no/sintef-xmlui/bitstream/handle/11250/2500523/CHI18_SIG_Chatbots%2Bfor%2Bsocial%2Bgood_authors%2Bversion.pdf?sequence=4 Date of access: 17 Mar. 2022.

Fonseka, T.M., Bhat, V. & Kennedy, S.H. 2019. The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. *Australian & New Zealand Journal of Psychiatry*, 53(10):954-964.

Foster, I., Ghani, R., Jarmin, R.S., Kreuter, F. & Lane, J. 2020. *Big Data and social science: Data science methods and tools for research and practice*. CRC Press.

Fouché, C.B. & Strydom, H. 2021. Mixed methods research. In: Fouche', C.B., Stryfom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 419-459.

Fulmer, R., Joerin, A., Gentile, B., Lakerink, L. & Rauws, M. 2018. Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR mental health*, 5(4), e9782. https://mental.jmir.org/2018/4/e64/

Fusch, P.I. and Ness, L.R. 2015. Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report,* 20(9): 1408-1416.

Galderisi, S., Heinz, A., Kastrup, M., Beezhold, J. & Sartorius, N. 2015. Toward a new definition of mental health. *World psychiatry*, 14(2):231-233.

Gandhi, A., Adhvaryu, K., Poria, S., Cambria, E. & Hussain, A., 2022. Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion*, 91:424-444.

Garcia, E. 2024. Deep Reinforcement Learning for Adaptive Drug Dosage Optimization: Utilizes deep reinforcement learning to optimize drug dosage based on patient response. Journal of AI-Assisted Scientific Discovery, 4(1):64-77.

Gibson, K. & Cartwright, C. 2013. Agency in young clients' narratives of counseling: 'It's whatever you want to make of it'. *Journal of Counseling Psychology*, 60(3):340–352.

Gibson, K. & Cartwright, C., 2014. Young people's experiences of mobile phone text counselling: Balancing connection and control. *Children and youth services review*, 43:96-104.

Giglio, F. 2021. Fintech: A literature review. *European Research Studies Journal*, 24(2B):600-627.

Giguère-Rancourt, A., Plourde, M., Racine, E., Couture, M., Langlois, M., … Simard, M. 2022. Goal management training and psychoeducation / mindfulness for treatment of executive dysfunction in Parkinson's disease: A feasibility pilot trial. *PLoS ONE*. 17(2):e0263108.

Gilat, I., & Shahar, G. 2007. Emotional first aid for a suicide crisis: comparison between Telephonic hotline and internet. *Psychiatry: Interpersonal and Biological Processes*, 70(1):12-18.

Glen, S. 2022. *What are Probability and Statistics?* https://www.statisticshowto.com/ Date of access: 18 Mar. 2022.

Global Legal Insights. 2023. *AI, Machine Learning & Big Data Laws and Regulations 2023 | South Africa.* https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/south-africa Date of access: 27 Jan. 2024.

Go, E. & Sundar, S.S. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97:304-316.

Goldberg, S.B., Flemotomos, N., Martinez, V.R., Tanana, M.J., Kuo, P.B., Pace, B.T., … Atkins, D.C. 2020. Machine learning and natural language processing in psychotherapy research: alliance as example use case. *Journal of Counseling Psychology*, 67(4):438–448.

Goldkind, L. 2021. Social Work and Artificial Intelligence: Into the Matrix. *Social Work,* 66(4):372–374.

Goldkind, L. 2022. Social Work and Artificial Intelligence: Into the Matrix. *Social Work* (Provisionally Accepted). doi: 10.1093/sw/swab028

Googins, B. & Godfrey, J. 1987. *Occupational social work*. Englewood Cliffs, NJ: Prentice Hall.

Groenveld, I. F., Proper, K. I., Absalah, S., van der Beek, A. J. & van Mechelen, W. 2011. An individually based lifestyle intervention for workers at risk for cardiovascular disease: A process evaluation. *American Journal of Health Promotion*, 25:396–401.

Grover, F., Wu, H.D., Blanford, C., Holcomb, S. & Tidler, D. 2002. Computer-using patients want internet services from family physicians. *Journal of Family Practice*, 51:570–572

Grudin, J., & Jacques, R. 2019. Chatbots, humbots, and the quest for artificial general intelligence. *Conference proceedings*. 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland. New York: Association for Computing Machinery. pp. 1-11.

Guest, G., MacQueen, K.M. & Namey, E.E. 2019. *Introduction to Applied Thematic Analysis.* Thousand Oaks: SAGE Publications.

Gupta, A., Hathwar, D. & Vijayakumar, A., 2020. Introduction to AI chatbots. *International Journal of Engineering Research and Technology*, 9(7):255-258.

Güzel, B. 2021. Futurism in Social Work: Transhumanism, Gamification of Human Life, Homo Roboticus and Care. *Journal of Society & Social Work*, 32(2):589-606.

Haber, Y., Levkovich, I., Hadar-Shoval, D. & Elyoseph, Z. 2024. The artificial third: A broad view of the effects of introducing generative artificial intelligence on psychotherapy (Preprint). *JMIR Mental Health*. 11:e54781.

Hajjaji, Y., Boulila, W., Farah, I.R., Romdhani, I. & Hussain, A. 2021. Big data and IoT-based applications in smart environments: A systematic review. *Computer Science Review*, 39:1-39.

Hamilton, M. 1960. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry,* 23(1):56–62.

Hanley, T. & Reynolds, D.J. 2009. Counselling Psychology and the internet: A review of the quantitative research into online outcomes and alliances within text-based therapy. *Counselling Psychology Review*, 24(2):4-13.

Harari, Y.N. 2016. *Homo Deus: A Brief History of Tomorrow*. London: Vintage.

Harari, Y.N. 2019. *21 Lessons for the 21st Century*. London: Vintage.

Hardeniya, N.H., Perkins, J., Chopra, D., Joshi, N. & Mathur, I. 2016. *Natural Language Process: Python and NLTK*. Birmingham: Pact.

Hassan, S.U., Ahamed, J. & Ahmad, K. 2022. Analytics of machine learning-based algorithms for text classification. *Sustainable operations and computers*, 3:238-248.

Hasson, F. & Keeney, S. 2011. Enhancing rigour in the Delphi technique research. *Technological Forecasting & Social Change,* 78:1695-1704.

Hasson, F., Keeney, S. & McKenna, H. 2000. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4):1008-1015.

Heath, S., Liddle, J. & Wiles, J. 2019. The Challenges of Designing a Robot for a Satisfaction Survey: Surveying Humans Using a Social Robot. *International Journal of Social Robotics*, 12:519-533.

Heimler, E. 1990. *Heimler scale of social functioning Rev. IV*. London: Eugene Heimler Trust.

Hiland, E.B. 2021. *Therapy Tech: The Digital Transformation of Mental Healthcare*. Minneapolis, MN: University of Minnesota Press.

Hoffer, J. 2002. *Modern system analysis and design*. 4th ed. New Jersey: Pearson Prentice.

Holzmeyer, C., 2021. Beyond 'AI for Social Good'(AI4SG): social transformations—not tech-fixes—for health equity. *Interdisciplinary Science Reviews*, 46(1-2):94-125.

Howick, J., Morley, J. & Floridi, L. 2021. An Empathy Imitation Game: Empathy Turing Test for Care- and Chat-bots. *Minds and Machines,* 31:457-461, art. #21. doi:10.1007/s11023-021-09555-w

Hsu, C. & Sandford, B. 2007. The Delphi Technique: Making Sense of Consensus. Practical Assessment, Research & Evaluation, 12, 1-8.http://pareonline.net/pdf/v12n10.pdf

Hsu, J.Y.J & Hasmath, R. 2017. A maturing civil society in China? The role of knowledge and professionalisation in the development of N.G.O.s. *China Information*, 31(1):22-42, 2017.

Hudson, W.W. 1982. The clinical measurement package: A field manual. Chicago: Dorsey.

Hull, T.D., Malgaroli, M., Connolly, P.S., Feuerstein, S. & Simon, N.M. 2020. Two-way messaging therapy for depression and anxiety: longitudinal response trajectories. *BMC psychiatry*, 20(1):1-12.

Hurbans, R. 2020. *Artificial Intelligence Algorithms*. New York, NY: Manning Publications.

Hussain, M. & Manhas, J. 2016. Artificial intelligence for big data: Potential and relevance. *International Academy of Engineering and Medical Research*, 1(1):1-5.

Inkster, B., Sarda, S. & Subramanian, V. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR Mhealth Uhealth*, 6(11):3-18.

Jafari, H., Kassan, A., Reay, G., & Climie, E. A. 2022. Resilience in Refugee Children and Youth: A Critical Literature Review. *Canadian Psychology/Psychologie canadienne,* 3, art. #09. doi: 10.31703/glr.2018(III-I).09

Jain, M., Kumar, P., Kota, R. & Patel, S.N. 2018. *Evaluating and informing the design of chatbots*. Paper delivered at the 2018 Designing interactive systems conference, China, Hong Kong. http://dx.doi.org/10.1145/3196709.3196735 Date of access: 16 Dec. 2022.

Javaid, M., Haleem, A., Singh, R.P. & Suman, R. 2022. Artificial intelligence applications for industry 4.0: A literature-based study. *Journal of Industrial Integration and Management*, 7(1):83-111.

Jirava, P. 2004. System development life cycle. *Industrial Management & Data Systems,* 109(3): 338-356.

Johnson, A. 2023. ChatGPT In Schools: Here's Where It's Banned—And How It Could Potentially Help Students. *Forbes*, 18 Jan. https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students/?sh=7c237b7e6e2c Date of access: 28 Jan. 2023.

Johnson, E.T., Huck, G., Brooks, J. Moser, E. Blake, J. & Chan, F. 2015. Solution-Focused Brief Therapy. In: Chan, F., Berven, N.L. & Thomas, K.R., eds. *Counseling Theories and Techniques for Rehabilitation and Mental Health Professionals*. New York: Springer. pp. 49-69.

Johnson, R.B. & Gray, R. 2010. A history of philosophical and theoretical issues for mixed methods research. In: Tashakkori, A. & Teddlie, C., eds. *Handbook of mixed methods in social and behavioural research*. Thousand Oaks, CA: Sage. pp. 69-94.

Jones, C.L.E., Hancock, T., Kazandjian, B. & Voorhees, C.M. 2022. Engaging the Avatar: The effects of authenticity signals during chat-based service recoveries. *Journal of Business Research*, 144:703-716.

Joulin, A., Grave, E., Bojanowski, P. & Mikolov, T. 2016. Bag of tricks for efficient text classification. https://arxiv.org/abs/1607.01759 Date of access 31 Dec. 2022.

Jung, S. 2019. Semantic vector learning for natural language understanding. *Computer Speech & Language*, 56:130-145.

Kaniki, A.M. 2006. Doing an information search. In: Terre Blanche, M., Durrheim, K. & painter, D., eds. *Research in practice*. Cape Town: Cape Town University Press. pp. 18-33.

Kasi, V., Keil, M., Mathiassen, L., & Pedersen, K. 2008. The post mortem paradox: A Delphi study of I.T. specialist perceptions. *European Journal of Information Systems*, 17:62-78.

Keenan, M.J.S. 2019. *Advanced Positioning, Flow, and Sentiment Analysis in Commodity Markets : Bridging Fundamental and Technical Analysis*. Newark, NJ: John Wiley & Sons.

Kejriwal, M. 2022. Knowledge graphs: a practical review of the research landscape. Information, 13:161-178.

Kenny, M. C., & Mceachern, A. G. 2004. Telephone counseling: Are offices becoming obsolete? *Journal of Counseling & Development*, 82:199–203.

Kettner, P.M., Moroney, R.M. & Lawrence, L.M. 2008. *Designing and Managing Programs.* 3rd ed. Thousand Oaks: Sage.

Khan, I.A. and Kumari, D., 2021. The Role of Analysis Phase of SDLC for Small Scale Business Application-A Review. *International Journal Of Humanities, Engineering, Science And Management*, 2(01):63-75.

Khan, Z.F. & Alotaibi, R.A. 2020. Applications of Artificial Intelligence and Big Data Analytics in m-Health: A Healthcare System Perspective. *Journal of Healthcare Engineering*, 2020, art. #8894694. doi:10.1155/2020/8894694

Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Bhale, P. & Das, T. 2015. A study of today's A.I. through chatbots and rediscovery of machine intelligence. *International Journal of u- and e- Service, Science and Technology*, 8(7):277-284.

Khurana, D., Koli, A., Khatter, K., & Singh, S. 2022. Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 37, art.#14. doi:10.1007/s11042-022-13428-4

Kirsten, N. 2022. Big data application in A.I. development. [Telephonic interview]. 2 May, Krugersdorp.

Knott, A. 2023. Job sharing between human professionals and chatbots: How should 'handovers' happen?. *Journal of AI, Robotics & Workplace Automation*, 2(2):121-130.

Koulouri, T., Macredie, R.D. & Olakitan, D. 2022. Chatbots to support young adults' mental health: an exploratory study of acceptability. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 12(2):1-39.

Koutsouleris, N., Hauser, T.U., Skvortsova, V. and De Choudhury, M. 2022. From promise to practice: towards the realisation of AI-informed mental health care. *The Lancet Digital Health*, 4(11):829-840.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A. & Singh, I. 2019. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights*, 11:1-9.

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I. & NeurOx Young People's Advisory Group. 2019. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. *Biomedical informatics insights*, 11, p.1178222619829083. https://doi.org/10.1177/1178222619829083

Krishnav, D. 2022. *Basics of building an Artificial Intelligence Chatbot* [Blog post]. https://www.mygreatlearning.com/blog/basics-of-building-an-artificial-intelligence-chatbot/#textchatbot Date of access: 01 May. 2022.

Kruger, A. & Van Breda, A.D. 2001. Military social work in the South African National Defence Force. *Military Medicine,* 166(11):947-951.

Kumar, A. & Kumar, A. 2022. Human Sentiment Analysis on Social Media through Naïve Bayes Classifier. *Journal of Scientific Research*, 66(1):350-357.

Kumar, R. & Ali, M.M. 2020. A review on chatbot design and implementation techniques. *Int. J. Eng. Technol*, 7(11):2791-2800.

Kumar, V.M., Keerthana, A., Valliammai, S. & Vinithasri, V. 2016. Sanative Chatbot For Health Seekers. *International Journal Of Engineering And Computer Science,* 5(3):16022-16025.

Kupferberg, A., Bicks, L. & Hasler, G. 2016. Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews*, 69:313-332.

Landowska, A. 2018. Towards New Mappings between Emotion Representation Models. *Applied Sciences*, 8(2), art. 274. https://doi.org/10.3390/app8020274

Lee, C.T., Pan, L.Y. & Hsieh, S.H. 2022. Artificial intelligent chatbots as brand promoters: a two-stage structural equation modelling-artificial neural network approach. *Internet Research*, 32(4): 1329-1356.

Lee, E.E., Torous, J., De Choudhury, M., Depp, C.A., Graham, S.A., Kim, H.C., Paulus, M.P., Krystal, J.H. & Jeste, D.V. 2021. Artificial intelligence for mental health care: clinical applications, barriers, facilitators, and artificial wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 6(9):856-864.

Levinson, C. 2018. *The Definition of Lean Operations.* https://bizfluent.com/about-6666969-definition-lean-operations.html Date of access: 19 Mar. 2022.

Lexalytics. 2020. *The 7 Basic Functions of Text Analytics & Text Mining.* https://www.lexalytics.com/lexablog/text-analytics-functions-explained Date of access: 30 Apr. 2022.

Li, S., Long, F., Liu, S., Lee, Q., Guo, R., Li, C. & Liu, Y. 2006. Herbs for medicinal baths among the traditional Yao communities of China. *Journal of ethnopharmacology*, 108:59-67.

Lim, S.M., Shiau, C.W.C., Cheng, L.J. & Lau, Y. 2022. Chatbot-delivered psychotherapy for adults with depressive and anxiety symptoms: a systematic review and meta-regression. *Behavior Therapy*, 53(2):334-347.

Linstone, A.H. & Turoff, M. 1975. *The Delphi Method: Techniques and Applications*. Boston: Addison-Wesley.

Lipman, J. & Distler, R. 2023. Schools Shouldn't Ban Access to ChatGPT. *Time*, 10 Jan. https://time.com/6246574/schools-shouldnt-ban-access-to-chatgpt/ Date of access: 28 Jan. 2023.

Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. 2013. I can help you change! An empathic virtual agent delivers behavior change health interventions. *A.C.M. Transactions on Management Information Systems (TMIS)*, 4(4):19.

Lommatzsch, A. & Katins, J. 2019. An Information Retrieval-based Approach for Building Intuitive Chatbots for Large Knowledge Bases. *Lernen, Wissen, Daten, Analysen*, 49(2):343-352.

Lucey, B. & Dowling, M. 2023. ChatGPT: our study shows AI can produce academic papers good enough for journals – just as some ban it. *The Conversation*, 26 Jan. https://theconversation.com/chatgpt-our-study-shows-ai-can-produce-academic-papers-good-enough-for-journals-just-as-some-ban-it-197762 Date of access: 28 Jan. 2023.

Luger, E. & Sellen, A. 2016. *Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents*. Paper delivered at the 2016 CHI Conference on Human Factors in Computing Systems, USA, New York. doi:10.1145/2858036.2858288

Luther, A. 2017. The Entity Mapper A: Data Visualization Tool for Qualitative Research Methods. *Leonardo*, 50(3):268-271.

Lyons, S.T., Schweitzer, L. & Ng, E.S. 2015. Resilience in the modern career. *Career Development International*, 20(4):363-383.

Macnab-Holding, C. 2024. *Exploring employees' subjective experience of psychological empowerment after adopting Artificial Intelligence technology*. Stellenbosch: Stellenbosch University. (Dissertation – Masters).

Magogodi, P.P.S. 2017. *The Use of Online Text Based Technologies as a Medium for Employee Counselling: Perceptions of Online Counsellors*. Johannesburg: University of the Witwatersrand. (Dissertation – MA).

Mallen, M.J. & Vogel, L.V. 2005. Online Counseling: A Need for Discovery. *The Counseling Psychologist,* 33(6):910-921.

Marcelle, E.T., Nolting, L., Hinshaw, S.P. & Aguilera, A, 2019. Effectiveness of a multimodal digital psychotherapy platform for adult depression: a naturalistic feasibility study. *JMIR mHealth and uHealth*, 7(1), e10948. https://mhealth.jmir.org/2019/1/e10948

Marsch, L.A. 2021. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacol.* 46:191–196.

Martinengo, L., Lum, E. & Car, J. 2022. Evaluation of chatbot-delivered interventions for self-management of depression: content analysis. *Journal of affective disorders*, 319:598-607.

Mathekga, H.L. & Sekudu, J. 2019. Employee wellness: Preparation for the fourth industrial revolution. *International Journal of Innovative Studies in Sociology and Humanities*, 4(5):31-39.

Mathieu, S.L., Uddin, R., Brady, M., Batchelor, S., Ross, V., Spence, S.H., Watling, D. & Kõlves, K. 2021. Systematic review: The state of research into youth helplines. *Journal of the American Academy of Child & Adolescent Psychiatry*, 60(10):1190-1233.

Mawani, A. & Nderu, L., 2020. Towards an online empathy assisted counselling web application. *EAI Endorsed Transactions on Context-aware Systems and Applications*, 7(22), e4. https://doi-org.10.4108/eai.23-12-2020.167792

McCarthy, J. 1956. Measures of the value of information. *Proceedings of the National Academy of Sciences*, 42(9):654-655.

McCarthy, J., Minsky, M.L., Rochester, N., & Shannon, C.E. 1956. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.

McGee, V. 2022. *Virtual Therapy vs. In-Person Therapy.* https://www.psychology.org/resources/virtual-therapy-vs-in-person/ Date of access: 06 Apr. 2022.

Mead, D. & Mosely, L. 2001. The use of the Delphi as a research approach. *Nurse researcher*, 8(4):4-23.

Mehanna, Y.S. & Mahmuddin, M. 2021. The Effect of Pre-processing Techniques on the Accuracy of Sentiment Analysis Using Bag-of-Concepts Text Representation. *SN Computer Science*, 2:237-250.

Mehta, P. & Pandya, S. 2020. A review on sentiment analysis methodologies, practices and applications. *International Journal of Scientific and Technology Research*, 9(2):601-609.

Meyer, T.J., Miller, M.L., Metzger, R.L. & Borkovec, T.D. 1990. Development and validation of the Penn state worry questionnaire. *Behaviour research and therapy*, 28(6):487-495.

Michel, N.M., Rowa, K., Young, L. & McCabe, R.E. 2016. Emotional distress tolerance across anxiety disorders. *Journal of anxiety disorders*, 40:94-103.

Mierzwa, S., Souidi, S., Conroy, T., Abusyed, M., Watarai, H., & Allen, T. 2019. On the Potential, Feasibility, and Effectiveness of Chat Bots in Public Health Research Going Forward. *Online journal of public health informatics*, 11(2), e4. doi:10.5210/ojphi.v11i2.9998

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. & Dean, J. 2013. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26:1-8.

Mills, K.A. 2019. *Big Data for Qualitative Research*. New York, NY: Routledge.

Mindell, D.A. & Reynolds, E. 2023. *The work of the future: building better jobs in an age of intelligent machines*. Cambridge, MA: MIT Press.

Molano, J.I.R., Lovelle, J.M.C., Montenegro, C.E., Granados, J.J.R. & Crespo, R.G. 2018. Metamodel for integration of internet of things, social networks, the cloud and industry 4.0. *Journal of ambient intelligence and humanized computing*, 9:709-723.

Molnar, A. & de Shazer, S. 1987. Solution-focused therapy: Toward the identification of therapeutic tasks. *Journal of Marital and Family Therapy*, 13(4):349-358.

Moor, J. 2006. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. AI Magazine, 27:87-91.

Moore, M.M. & Ophir, Y. 2021. Big Data, Actually: Examining Systematic Messaging in 188 Romantic Comedies Using Unsupervised Machine Learning. *Psychology of Popular Media*, 14(3), art. #349. doi:10.1037/ppm0000349

Morrissey, D. 2021. *Social Work and Helping Humanity Through Chatbots*. https://medium.com/@devinmorrissey.1987/social-work-and-helping-humanity-through-chatbots-a669dab63b4d Date of access: 14 Jan. 2023.

Morrissey, D. 2021. *Social Work and Helping Humanity Through Chatbots*. https://medium.com/@devinmorrissey.1987/social-work-and-helping-humanity-through-chatbots-a669dab63b4d Date of access: 14 Jan. 2023.

Mostert-Phipps, N. Pottas,D. & Korpela, M. 2013. A South African perspective on factors that impact the adoption and meaningful use of health information technologies. *S Afr Fam Pract,* 55(6):545-554.

Munn, Z., Peters, M.D.J., Stern, C., Tufanaru, C., McArthur, A. & Aromataris, E. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *B.M.C. Med Res Methodol*, 18:143.

Nagarajan, M. & Yuvaraj, S. 2021. Mental health counsellors' perceptions of the use of technology in counselling. *Current Psychology*, 40:1760-1766, art. #4. doi:10.1007/s12144-018-0104-4

Nandwani, P. & Verma, R. 2021. A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1):81-100.

Ndzendze, B., Singh, A., Timm, S., Meyers, C., Malatji, M.J., Malatji, M.J., … Brokensha, S.I. 2024. *The 4IR and the Humanities in South Africa: Perspectives on innovation, power and potentialities*. Cape Town: AVARSITY Books.

Neuman, W. 2003. Social Research Methods: Qualitative and Quantitative Approaches. Essex: Pearson.

Nolan, B. 2023. A freelance writer says ChatGPT wrote a R10,000 article in just 30 seconds. *Business Insider South Africa*, 25 Jan. https://www.businessinsider.co.za/chatgpt-wrote-600-dollar-article-30-seconds-freelance-writer-ai-2023-1 Date of access: 28 Jan. 2023.

Nosrati, S., Sabzali, M., Heidari, A., Sarfi, T. & Sabbar, S. 2020. Chatbots, counselling, and discontents of the digital life. *Journal of Cyberspace Studies*, 4(2):153-172.

Novotney, A. 2017. A growing wave of online therapy. *Monitor on Psychology*, 48(2):48.

Noyce, R. & Simpson, J. 2016. The Experience of Forming a Therapeutic Relationship from the Client's Perspective: A Metasynthesis. *Psychotherapy Research*. 28(2):281-296.

Ogbeifun, E., Agwa-Ejon, J., Mbohwa, C. & Pretorius, J.H.C. 2016. The Delphi technique: A credible research methodology. *Industrial Engineering and Operations Management,* 8-10.

Oh, K.J., Lee, D., Ko, B. & Choi, H.J. 2017. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. Paper delivered at the 18th IEEE international conference on mobile data management (MDM), Daejeon. https://ieeexplore.ieee.org/document/7962482 Date of access: 02 Dec. 2023.

OpenAI. 2023. Join us in shaping the future of technology. https://openai.com/ Date of access: 28 Jan. 2023.

Pandian, A.P. 2021. Performance evaluation and comparison using deep learning techniques in sentiment analysis. *Journal of Soft Computing Paradigm*, 3(2):123-134.

Patel, K.K., Patel, S.M. & Scholar, P. 2016. Internet of things-IOT: definition, characteristics, architecture, enabling technologies, application & future challenges. *International journal of engineering science and computing*, 6(5):6122-6131.

Patterson, D.A. & Cloud, R.N. 1999. The application of artificial neural networks for outcome prediction in a cohort of severely mentally ill outpatients. *Journal of Technology for Human Services*, 16(3):47-61.

Paul, J. & Criado, A.R. 2020. The art of writing literature review: What do we know and what do we need to know? International Business Review, 29(4):101-108.

Pencheva, I., Esteve, M. and Mikhaylov, S.J. 2020. Big Data and AI–A transformational shift for government: So, what next for research?. *Public Policy and Administration*, 35(1):24-44.

Pinciroli, F., Justo, J.L.B. & Forradellas, R. 2022. Systematic mapping study: On the coverage of aspect-oriented methodologies for the early phases of the software development life cycle. *Journal of King Saud University-Computer and Information Sciences*, 34(6):2883-2896.

Pinciroli, F., Justo, J.L.B. & Forradellas, R. 2022. Systematic mapping study: On the coverage of aspect-oriented methodologies for the early phases of the software development life cycle. *Journal of King Saud University-Computer and Information Sciences*, 34(6):2883-2896.

Pink, S., Ferguson, H. & Kelly, L. 2022. Digital social work: Conceptualising a hybrid anticipatory practice. *Qualitative Social Work*, 21(2):413-430.

Poh Li, L., Rafidah, A.M.J. & Abdullah, H.S. 2013. Understanding the two sides of online counseling and their ethical and legal ramifications. *Procedia - Social and Behavioral Sciences*, 103:1243-1251.

Porter, N.D., Verdery, A.M. & Gaddis, S.M. 2020. Enhancing big data in the social sciences with crowdsourcing: Data augmentation practices, techniques, and opportunities. *PLoS ONE*, 15(6), e0233154. https://doi.org/10.1371/journal. pone.0233154

Pothukuchi, A.S., Kota, L.V. and Mallikarjunaradhya, V. 2023. Impact of Generative AI on the Software Development Lifecycle (SDLC). *International Journal of Creative Research Thoughts*, 11(8):227-291.

Potts, C., Ennis, E., Bond, R.B., Mulvenna, M.D., McTear, M.F., … O'Neill, S. 2021. Chatbots to support mental wellbeing of people living in rural areas: Can user groups contribute to co-design? *Journal of Technology in Behavioral Science.* 6(4):652–665.

Pradhan, A., Senapati, M. R. & Sahu, P. K. 2022. Improving Sentiment Analysis with Learning Concepts from Concept, Patterns Lexicons and Negations. *Ain Shams Engineering Journal*, 13(2), art. #101559. doi:10.1016/j.asej.2021.08.004

Prochaska, J.J., Vogel, E.A., Chieng, A., Kendra, M., Baiocchi, M., Pajarito, S. & Robinson, A. 2021. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. *Journal of medical Internet research*, 23(3), e24850. doi:10.2196/24850

Rathnayaka, P., Mills, N., Burnett, D., De Silva, D., Alahakoon, D. & Gray, R. 2022. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors*, 22(10):3653-3668.

Ray, P. & Chakrabarti, A. 2017. Twitter sentiment analysis for product review using lexicon method. Paper delivered at the International Conference on Data Management, Analytics, and Innovation (ICDMAI), India, Pune. https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/apirference/get-search-tweets Date of access: 30 Apr. 2022.

Rice, E., Wilder, B., Onasch-Vera, L., DiGuiseppi, G., Petering, R., Hill, C., Yadav, A., Lee, S.J. & Tambe, M. 2021. A Peer-Led, Artificial Intelligence–Augmented Social Network Intervention to Prevent HIV Among Youth Experiencing Homelessness. *Journal of Acquired Immune Deficiency Syndromes*, 88(1):20-26.

Rivas, P., Holzmayer, K., Hernandez, C. & Grippaldi, C. 2018. Excitement and concerns about machine learning-based chatbots and talkbots: A survey. *Conference proceedings*. IEEE International Symposium on Technology and Society (ISTAS), Washinton, DC, United States of America. Institute of Electrical and Electronics Engineer. pp. 156-162.

Roestenburg, W.J.H. 2021. *Analysing technology acceptance by social workers in South Africa*. Potchefstroom: North-West University. (Dissertation – M.B.A.).

Rothman, J. & Thomas, E.J. 1994. *Intervention Research: Design and Development for Human Service*. New York: Haworth Press.

Rowe, G., Wright, G. & McColl, A. 2005. Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence. *Technology Forecast and Social Change*, 72:377- 399.

Royce, W.W. 1970. Managing the Development of Large Software Systems. *Conference proceedings.* IEEE WESCON. Los Angeles, United States: Western Electronic Show and Convention. pp. 328-388.

Rubin, A. & Babbie, E. 2013. *Essential Research Methods for Social Work*. 3rd ed. Boston, MA: Cengage Learning.

Rudestam, K.E. & Newton, R.R. 1992. *Surviving your dissertation: A comprehensive guide to content and process*. California: Sage.

Russel, S.J. & Norvig, P. 2022. *Artificial Intelligence: A Modern Approach*. 4th ed. Harlow: Pearson Education.

Safodien, M. 2021. Social work 4.0? The fourth industrial revolution and social work education: a South African perspective. *Social Work/Maatskaplike Werk*, 57(3):257-271.

Sage, S. & Singer, J.B. 2022. Technology and Social Work Practice. Micro, Mezzo and Macro Applications. In: Rapp-McCall, L., Roberts, A. & Corcoran, K., eds. *Social Workers' Desk Reference*. New York, NY: Oxford University Press. pp. 45-52.

Sakurai, E., Kurashige, K., Tsuruta, S., Sakurai, Y., Knauf, R., Damiani, E., … Frati, F. 2020. Embodiment matters: toward culture-specific robotized counselling. *Journal of Reliable Intelligent Environments*, 6:129-139.

Saliba, Y. & Barden, S. 2017. Counselors and Workplace Wellness Programs: A Conceptual Model. *The Professional Counselor*, 7(2):104-113.

Sánchez, P.M., Aburto, L.G., Orellana, M.O. & Torres, R.C. 2020. Knowledge and skills for occupational social work: challenges for education and professional practice. *Cuadernos de Trabajo Social*, 33(1):247-258.

Sankar, R. 2018. Empowering chatbots with business intelligence by big data integration. *International Journal of Advanced Research in Computer Science*, 9(1):627-631.

Santos, G.A., De Andrade, G.G., Silva, G.R.S., Duarte, F.C.M., Da Costa, J.P.J. & De Sousa, R.T. 2022. A Conversation-Driven Approach for Chatbot Management. *IEEE Access, 10:8474-8486. doi:*10.1109/ACCESS.2022.3143323

Santos, M.Y. & Costa, C. 2020. *Big Data: Concepts, Warehousing, and Analytics*. New York, NY: Routledge.

Saravanos, A. & Curinga, M.X. 2023. Simulating the Software Development Lifecycle: The Waterfall Model. *Applied System Innovation*, 6(6):108-127.

Sarker, I.H., Khan, A.I., Abushark, Y.B. & Alsolami, F. 2021. Mobile expert system: exploring context-aware machine learning rules for personalized decision-making in mobile applications. *Symmetry*, 13(10):1975-1985.

Sauer, P.C. & Seuring, S. 2023. How to conduct systematic literature reviews in management research: a guide in 6 steps and 14 decisions. Review of Managerial Science, 17(5):1899-1933.

Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R. & Grant, M. 1993. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, 88(6):791-804.

Schlenger, W., Jöllenbeck, M., Stamer, T., Grosse, A. & Ochsmann, E. 2022. Digitising Social Counseling—Insights for Workplace Health Management. *International Journal of Environmental Research and Public Health*, 19:1-16.

Schnurr, P., Vielhauer, M., Weathers, F. & Findler, M. 1999. Brief trauma questionnaire. *APA PsycTests*. https://doi.org/10.1037/t07488-000

Schuetzlera, R.M., Mark, G. & Giboneya, J.S. 2020. Impact of Chatbot Conversational Skill on Engagement and Perceived Humanness. *Journal of management information systems,* 37(3): 875-900.

Schurink, W.J., Schurink, E.M. & Fouché, C.B. 2021. Qualitative data analysis and interpretation. In: Fouche', C.B., Stryfom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 391-415.

Schwab, K. 2015. *The Fourth Industrial Revolution: what it means, how to respond.* https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/ Date of access: 11 Mar. 2022.

Schwab, K. 2016. *The Fourth Industrial Revolution*. Geneva: World Economic Forum.

Selzer, M.L., Vinokur, A. & van Rooijen, L. 1975. A self-administered short Michigan alcoholism screening test (SMAST). *Journal of studies on alcohol*, 36(1):117-126.

Sepahpour, T. 2020. *Ethical considerations of chatbot use for mental health support*. Baltimore: Johns Hopkins University. (Thesis – Masters).

Setyawan, M.Y.H., Awangga, R.M. & Efendi, S.R. 2018. Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot. *Conference proceedings*.

2018 International Conference on Applied Engineering (ICAE, 2018), Batam, Indonesia. Institute of Electrical and Electronics Engineers. pp. 1-5.

Setyawan, M.Y.H., Awangga, R.M. & Efendi, S.R. 2018. Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot. *Conference proceedings*. 2018 International Conference on Applied Engineering (ICAE, 2018), Batam, Indonesia. Institute of Electrical and Electronics Engineers. pp. 1-5.

Shah, S.R., Ahmed, F. & Khan, R. 2018. Writing a Critical Review of Literature: A Practical Guide for English Graduate Students. *Global Language Review*, 3:136-153.

Sharma, M.K., 2017. A study of SDLC to develop well engineered software. *International Journal of Advanced Research in Computer Science*, 8(3):520-523.

Shinozaki, T., Yamamoto, Y. and Tsuruta, S., 2015. Context-based counselor agent for software development ecosystem. *Computing*, 97:3-28.

Shreyas, J., Jumnal, A., Kumar, S.D. & Venugopal, K.R. 2020. Application of computational intelligence techniques for internet of things: an extensive survey. *International Journal of Computational Intelligence Studies*, 9(3):234-288.

Shum, H.Y., He, X.D. & Li, D. 2018. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*:10-26.

Singh, A., Ramasubramanian, K. & Shivam, S. 2019. Building an enterprise chatbot: work with protected enterprise data using open source frameworks. New York, NY: Apress.

Sinha, S.K. & Dubey, A.D. 2022. Artificial Intelligence Concerning Corporate Sustainability: A Systematic Review of Research. In: Goel, R., Baral, S.K. & Venkatesh, R., eds. Internet of Things and Machine Learning. New York, NY: Nova Science Publishers. pp. 215-233.

Skinner, R., Nelson, R.R., Chin, W.W. & Land, L. 2015. The Delphi Method Research Strategy in Studies of Information Systems. *Communications of the Association for Information Systems,* 37:31-63.

Skjuve, M.B. & Brandtzaeg, P.B. 2018. Chatbots as a new user interface for providing health information to young people. https://sintef.brage.unit.no/sintef-xmlui/ Date of access: 05 Oct. 2021.

Slomski, A., 2015. Online therapy reduces suicide ideation in medical interns. *Jama*, 314(24):2608-2618.

Smys, S. & Haoxiang, W. 2021. Naïve Bayes and entropy based analysis and classification of humans and chat bots. *Journal of IoT in Social Mobile Analytics and Cloud*, 3(1):40-49.

Solanki, P., Grundy, J. & Hussain, W. 2023. Operationalising ethics in artificial intelligence for healthcare: A framework for AI developers. *AI and Ethics*, 3(1):223-240.

Sperandeo, R., Cioffi, V., Mosca, L.L., Longobardi, T., Moretto, E., Alfano, Y.M., … Maldonato, N.M. 2021. Exploring the Question: Does Empathy Work in the Same Way in Online and In-Person Therapeutic Settings? *Front Psychol.,* 21(12), art. #671790. doi: 10.3389/fpsyg.2021.671790

Spitzer, R.L., Kroenke, K., Williams, J.B. & Löwe, B. 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166(10):1092-1097.

Srinivasan, J. 2021. *2 Ways AI is Transforming Wellness in the Workplace*. https://www.spiceworks.com/hr/engagement-retention/guest-article/2-ways-ai-is-transforming-wellness-in-the-workplace/ Date of access: 14 Jan. 2023.

Srinivasan, J. 2021. *2 Ways AI is Transforming Wellness in the Workplace*. https://www.spiceworks.com/hr/engagement-retention/guest-article/2-ways-ai-is-transforming-wellness-in-the-workplace/ Date of access: 14 Jan. 2023.

Steidl, M., Felderer, M. Ramler, R. 2023. The pipeline for the continuous development of artificial intelligence models—Current state of research and practice. Journal of Systems and Software, 199:111615.

Strydom, H. 2021. Sampling techniques and pilot studies in qualitative research. In: Fouché, C.B., Strydom, H. & Roestenburg, W.J.H. eds*. Research at Grass Roots*. Pretoria: Van Schaik Publishers. pp. 379-390.

Sullivan, G.M. 2011. A primer on the validity of assessment instruments. *J Grad Med Educ*. 3:119-20.

Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J… 2019. *Mitigating gender bias in* natural language processing: Literature review. https://arxiv.org/abs/1906.08976 Date of access: 30 Mar. 2023.

Sun, T.X., Liu, X.Y., Qiu, X.P. & Huang, X.J. 2022. Paradigm shift in natural language processing. *Machine Intelligence Research*, 19(3):169-183.

Sunter, C. & Ilbury, M. 2021. *Thinking the future: New Perspectives from the shoulders of giants.* Cape Town: Penguin Books.

Susanti, A.R., Djanta, T.& Kusuma, W.A. 2017. Twitter's Sentiment Analysis on Gsm Services using Multinomial Naïve Bayes. *Telkomnika,* 15(3):1354-1361.

Suta, P., Lan, X., Wu, B., Mongkolnam, P. & Chan, J.H. 2020. An overview of machine learning in chatbots. *International Journal of Mechanical Engineering and Robotics Research*, 9(4):502-510.

Taecharungroj, V. 2023. “What Can ChatGPT Do?” Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, 7(1):35.

Tambe, M. & Rice, E. 2018. Merging Social Work Science and Computer Science for Social Good. In: Tambe, M & Rice, E., eds. *Artificial Intelligence and Social Work*. Cambridge: Sheridan Books. pp. 3-16.

Tambe, M. & Rice, E. 2018. Merging Social Work Science and Computer Science for Social Good. In: Tambe, M & Rice, E., eds. *Artificial Intelligence and Social Work*. Cambridge: Sheridan Books. pp. 48-72.

Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T. and Qin, B. 2014. Learning sentiment-specific word embedding for twitter sentiment classification. *Conference proceedings.* 52nd Annual Meeting of the Association for Computational Linguistics (ACL, 2014), Baltimore, United States of America. Baltimore: Association for Computational Linguistics. pp. 1555-1565.

Tebenkov, E. and Prokhorov, I. 2021. Machine learning algorithms for teaching AI chat bots. *Procedia Computer Science*, 190:735-744.

Terblanche, L. 2021. Employee assistance programmes and occupational social work: interrelated with distinct features. *Social Work/Maatskaplike Werk*, 57(1):16-38.

Terrell, S. R. 2012. Mixed-Methods Research Methodologies. *The Qualitative Report*, *17*(1):254-280.

Tongco, C.D. 2007. Purposive sampling as a tool for informant selection. *Ethnobotany research & applications,* 5:147-158.

Topooco, N., Berg, M., Johansson, S., Liljethörn, L., Radvogin, E., Vlaescu, G… Andersson, G. 2018. Chat-and internet-based cognitive–behavioural therapy in treatment of adolescent depression: randomised controlled trial. *BJPsych open*, 4(4):199-207.

Topooco, N., Byléhn, S., Dahlström, E., Holmlund, J., Lindegaard, J… Andersson, G. 2019. Evaluating the efficacy of internet-delivered cognitive behavioral therapy blended with synchronous chat sessions to treat adolescent depression: randomized controlled trial. *Journal of medical Internet research*, 21(11), e13393. https://www.jmir.org/2019/11/e13393

Toyama, M., Fuller, H.R. & Hektner, J.M. 2020. Psychosocial Factors Promoting Personal Growth Throughout Adulthood. *Journal of Happiness Studies,* 21:1749-1769.

Trappey, A.J., Lin, A.P., Hsu, K.Y., Trappey, C.V. & Tu, K.L., 2022. Development of an empathy-centric counseling chatbot system capable of sentimental dialogue analysis. *Processes*, 10(5), 930. https://doi.org/10.3390/pr10050930

Turing, A.M. 1950. Computing Machinery and Intelligence. *Mind,* 59(236):433–60.

Tutun, S., Johnson, M.E., Ahmed, A., Albizri, A., Irgil, S., Yesilkaya, I,… 2023. An AI-based decision support system for predicting mental health disorders. *Information Systems Frontiers*, 25(3):1261-1276.

Umoquit, M., Tso, P., Varga-Atkins, T., O'Brien, M., & Wheeldon, J. 2013. Diagrammatic elicitation: Defining the use of diagrams in data collection. *The Qualitative Report*, 18(30):1-7.

Vaidyam, A.N., Wisniewski, H., Halamka, J.D., Kashavan, M.S. & Torous, J.B. 2019. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7):456-464.

Valério, A.M., Guimarães, T.G., Prates, R.O. & Candello, H. 2017. *Here's What I Can Do: Chatbots' Strategies to Convey Their Features to Users.* Paper delivered at the XVI Brazilian Symposium on Human Factors in Computing Systems, Joinville. https://dl.acm.org/doi/10.1145/3160504.3160544 Date of access: 18 Dec. 2022.

Van Ewijk, H. 2010. *European social policy and social work. Citizen-based social work*. Abingdon: Routledge.

Van Ewijk, H. 2018. *Complexity and social work.* Abingdon: Routledge.

Venkatesh, V., Brown, S.A. & Sullivan, Y.W. 2016. Guidelines for Conducting Mixed-methods Research: An Extension and Illustration. *Journal of the Association for Information Systems*, 17(7):436-495.

Vijayaraghavan, V., Cooper, J.B. & Leevinson, R.J. 2020. Algorithm inspection for chatbot performance evaluation. *Procedia Computer Science*, 171:2267-2274.

Vilaza, G.N. & McCashin, D. 2021. Is the Automation of Digital Mental Health Ethical? Applying an Ethical Framework to Chatbots for Cognitive Behaviour Therapy. *Frontiers in Digital Health*, 3, art. #2673. doi:10.3389/fdgth.2021.689736

Visvikis, D., Le Rest, C.C., Jaouen, V. & Hatt, M. 2019. Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications. *European Journal of Nuclear Medicine and Molecular Imaging*, 46:2630-2637.

Wahyono, I.D., Fadlika, I., Asfani, K., Putranto, H. & Hammad, J. 2019. New adaptive intelligence method for personalised adaptive laboratories. *Electrical, Electronics and Information Engineering,* 6:196-200.

Wallace, R.S. 2009. The anatomy of ALICE. In: Epstein, R. & Peters, G., eds. *Parsing the Turing test*. Dordrecht: Springer. pp. 181-210.

Walliman, N. 2011. *Research Methods the basics*. London: Routledge.

Walter, D. & Ophir, Y. 2019. News Frame Analysis: An Inductive Mixed-method Computational Approach. *Communication Methods and Measures*, 13(4):248-266.

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., … Zhang, W., 2022. A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion*, 83:19-52.

Wang, Z. 2023. Contextual Embedding Representations for Retrieval-Based and Generation-Based Dialogue Systems. Atlanta: Emory University. (Thesis - PhD).

Watson, D. & Friend, R. 1969. Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology,* 33(4):448–457.

Weizenbaum, J. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1):36-45.

Wu, E.H.K., Lin, C.H., Ou, Y.Y., Liu, C.Z., Wang, W.K. & Chao, C.Y. 2020. Advantages and constraints of a hybrid model K-12 E-Learning assistant chatbot. *IEEE Access,* 8:77788-77801. doi:10.1109/ACCESS.2020.2988252

Wulf, A. 2020. How to use text messages as a dialogue tool in Solution Focused conversations. In: Dierolf, K., Hogan, D. & van der Hoorn, S., eds. *Solution Focused Practice Around the World*. New York: Routledge. pp. 94-107.

Xu, M. & David, J.M. & Kim, S.H. 2018. The Fourth Industrial Revolution: Opportunities and Challenges. *International Journal of Financial Research,* 9(2):90-95.

Xu, S., Li, Y. & Zheng, W. 2017. Bayesian Multinomial Naïve Bayes Classifier to Text Classification. Lecture Notes in Electrical Engineering*,* 448:347-352.

Yadav, A. & Vishwakarma, D.K., 2020. Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review*, 53(6):4335-4385.

Yang, H., Lee, Y.I., Lee, H.J., Cho, S.W. & Koo, M.W. 2015. A study on word vector models for representing Korean semantic information. *Phonetics and Speech Sciences,* 7(4):41-47.

Yarowsky, D. 1995. Unsupervised word sense disambiguation rivalling supervised methods. Paper delivered at the 33rd annual meeting of the association for computational linguistics, Cambridge, MA. https://aclanthology.org/P95-1000/ Date of access: 14 Jan. 2023.

Yates, T.M. & Frank, L.S. 2021. Natural language processing: Opportunities and challenges for social work research. *Social Work Research*, 45(3):145-156.

Yildirim, S. & Asgari-Chenaghlu, M. 2021. Mastering Transformers: Build state-of-the-art models from scratch with advanced natural language processing techniques. Packt: Birmingham.

Zetino, J. & Mendoza, N. 2019. Big Data and Its Utility in Social Work: Learning from the Big Data Revolution in Business and Healthcare. *Social Work in Public Health*, 34(5):409-417.

Zhang, Q. 2022. A Big Data-Driven Approach to Analyse the Influencing Factors of Enterprise's Technological Innovation. *Computational Intelligence and Neuroscience*, 2022, art. #3785685. doi:10.1155/2022/3785685

Zhu, Y., Janssen, M., Wang, R. & Liu, Y. 2022. It is me, chatbot: working to address the COVID-19 outbreak-related mental health issues in China. User experience, satisfaction, and influencing factors. *International Journal of Human–Computer Interaction*, 38(12):1182-1194.

Zumstein, D. & Hundertmark, S. 2017. Chatbots - An Interactive Technology for Personalized Communication, Transactions and Services. *IADIS International journal on www/internet*, 15(1):96-109.

ANNEXURE 1: SCIENTIFIC COMMITTEE APPROVAL

Private Bag X1290, Potchefstroom
South Africa 2520

Tel: 086 016 9698
Web: http://www.nwu.ac.za/

North-West University Health Research Ethics Committee (NWU-HREC)

Tel: 018 299-1206
Email: Ethics-HRECApply@nwu.ac.za (for human studies)

29 November 2022

ETHICS APPROVAL LETTER OF STUDY

Based on approval by the North-West University Health Research Ethics Committee (NWU-HREC) on 29/11/2022, the NWU-HREC hereby approves your study as indicated below. This implies that the NWU-HREC grants its permission that, provided the general and specific conditions specified below are met and pending any other authorisation that may be necessary, the study may be initiated, using the ethics number below.

Study title: Developing an intelligent online counselling framework for an employee health and wellness programme
Principal Investigator/Study Supervisor/Researcher: Prof WJH Roestenburg
Student: I Kanes - 20728670

Ethics number: N W U - 0 0 1 7 8 - 2 2 - A 1
Institution Study Number Year Status

Status: S = Submission; R = Re-Submission; P = Provisional Authorisation; A = Authorisation

Application Type: Single study
Commencement date: 29/11/2022
Expiry date: 30/11/2023

Risk: **Medium**

Approval of the study is provided for a year, after which continuation of the study is dependent on receipt and review of a six-monthly monitoring report and the concomitant issuing of a letter of continuation. Monitoring reports are due at the end of May and November annually until completion of the study.

General conditions:

While this ethics approval is subject to all declarations, undertakings and agreements incorporated and signed in the application form, the following general terms and conditions will apply:

- *The principal investigator/study supervisor/researcher must report in the prescribed format to the NWU-HREC:*
  - *six-monthly the monitoring of the study, whereby a letter of continuation will be provided annually, and upon completion of the study; and*
  - *without any delay in case of any adverse event or incident (or any matter that interrupts sound ethical principles) during the course of the study.*
- *The approval applies strictly to the proposal as stipulated in the application form. Should any amendments to the proposal be deemed necessary during the course of the study, the principal investigator/study supervisor/researcher must apply for approval of these amendments at the NWU-HREC, prior to implementation. Should there be any deviations from the study proposal without the necessary approval of such amendments, the ethics approval is immediately and automatically forfeited.*
- *Annually a number of studies may be randomly selected for active monitoring.*
- *The date of approval indicates the first date that the study may be started.*
- *In the interest of ethical responsibility, the NWU-HREC reserves the right to:*

- *request access to any information or data at any time during the course or after completion of the study;*
- *to ask further questions, seek additional information, require further modification or monitor the conduct of your research or the informed consent process;*
- *withdraw or postpone approval if:*
  - *any unethical principles or practices of the study are revealed or suspected;*
  - *it becomes apparent that any relevant information was withheld from the NWU-HREC or that information has been false or misrepresented;*
  - *submission of the six-monthly monitoring report, the required amendments, or reporting of adverse events or incidents was not done in a timely manner and accurately; and/or*
  - *new institutional rules, national legislation or international conventions deem it necessary.*

- *NWU-HREC can be contacted for further information via Ethics-HRECApply@nwu.ac.za or 018 299 1206*

Special conditions of the research approval due to the COVID-19 pandemic:

Please note: Due to the nature of the study i.e. (online collection of both quantitative and qualitative data via document analysis of live chat transcript data and an online Delphi study with subject experts), this study will be able to proceed during the current alert level, following receipt of the approval letter. No additional COVID-19 restrictions have been placed on the study except that the researcher must ensure that before proceeding with the study that all research team members have reviewed the North-West University COVID-19 Occupational Health and Safety Standard Operating Procedure.

The NWU-HREC would like to remain at your service and wishes you well with your study. Please do not hesitate to contact the NWU-HREC for any further enquiries or requests for assistance.

Yours sincerely,

Digitally signed by
Prof Petra Bester
Date: 2022.11.30
14:30:51 +02'00'

Chairperson NWU-HREC

Current details (23239522) G:\My Drive\9. Research and Postgraduate Education\9.1.5.4 Templates\9.1.5.4.2_NWU-HREC_EAL.docm
20 August 2019
File Reference: 9.1.5.4.2

ANNEXURE 2: LETTER OF AUTHORISATION TO CONDUCT RESEARCH AT ICAS SOUTHERN AFRICA

ICAS South Africa
A Lyra Health company

5th Floor, Sandton Gate
7 Minerva Ave
Glenadrienne
Sandton

Postal:
PO Box 522822
Saxonwold, 2132

+27 (0) 11 380 6800
+27 (0) 11 325 7122
info@icas.co.za
www.icas.co.za

Health Research Ethics Committee

North-West University
Potchefstroom Campus
Private Bag X6001
Potchefstroom
2520

Letter of Authorization to Conduct Research at ICAS Southern Africa.

The NWU Heath Research Ethics Committee:

This letter will serve as authorisation for the NWU Potchefstroom researcher, Ivan Kanes to conduct the research project entitled Developing an Intelligent Online Counselling App for an Employee Health and Wellness Programme at ICAS Southern Africa (the company).

ICAS Southern African acknowledges that it has reviewed the protocol presented by the researcher, as well as the associated risks to the company. The company accepts the protocol and the associated risks to the company, and authorises the research project to proceed. The research project may be implemented at the company upon approval from the NWU Health Research Ethics Committee.

If we have any concerns or require additional information, we will contact the researcher and/or the NWU Health Research Ethics Committee.

Sincerely,

C Holland

Chantell Holland

Clinical and Operations Director

ICAS Southern Africa

18 August 2022

Date

ANNEXURE 3: LETTER OF AUTHORISATION TO CONDUCT RESEARCH AT LYRA SA – ACCESS TO AFFILIATE NETWORK

lyra | Southern Africa
Formerly ICAS Southern Africa

5th Floor, Sandton Gate
7 Minerva Ave
Glenadrienne Sandton

Postal:
PO Box 522822
Saxonwold, 2132

+27 (0) 11 380 6800
+27 (0) 11 325 7122
info@icas.co.za
www.icas.co.za

Health Research Ethics Committee
North-West University
Potchefstroom Campus
Private Bag X6001
Potchefstroom
2520

Letter of Authorization to Conduct Research at Lyra Southern Africa

This letter serves as authorization for the NWU Potchefstroom researcher, Ivan Kanes, to conduct the research project entitled **Developing an intelligent online counselling framework for an employee health and wellness programme** at Lyra Southern Africa.

Lyra Southern Africa acknowledges that it has reviewed the protocol presented by the researcher, as well as the associated risks to the company. The company accepts the protocol and the associated risks to the company and authorizes the research project to proceed. The research project may be implemented at the company, as approval from the NWU Health Research Ethics Committee has been granted.

Additionally, Lyra Southern Africa grants access to the Lyra Southern Africa clinical affiliate network to facilitate a Delphi panel discussion in the form of a focus group with clinical affiliates as part of the research project.

If we have any concerns or require additional information, we will contact the researcher and/or the NWU Health Research Ethics Committee.

Sincerely,



15 July 2024

Lauren Fitchet-Hall
Clinical Head
Lyra Southern Africa

Date

REGIONAL OFFICES:
Kwazulu-Natal +27 (0) 31 764 4882
Western Cape +27 (0) 21 673 6500

Lyra Southern Africa (Pty) Ltd Registered in South Africa No. 1998/008545/07 VAT No.4760174203

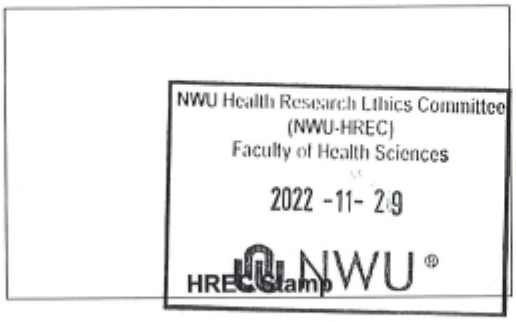
Part of the group of Lyra Health Holdings, LLC, a company incorporated under the laws of Delaware with file number 6410815, having its registered office at 251 Little Falls Drive, Wilmington, DE 19808

This information is classified as General (INTERNAL) and may not be shared with external parties.

ANNEXURE 4: INFORMED CONSENT DOCUMENTATION FOR ICAS LIVE CHAT COUNSELLORS

NWU®
NORTH-WEST UNIVERSITY
NOORDWES-UNIVERSITEIT
YUNIBESITI YA BOKONE-BOPHIRIMA

Private Bag X1290, Potchefstroom
South Africa 2520
Tel: +2718 299-1111/2222
Fax: +2718 299-4910
Web:http://www.nwu.ac.za



INFORMED CONSENT DOCUMENTATION FOR: ICAS LIVE CHAT COUNSELLORS

TITLE OF THE RESEARCH STUDY: DEVELOPING AN INTELLIGENT ONLINE COUNSELLING APP FOR AN EMPLOYEE HEALTH AND WELLNESS PROGRAMME

ETHICS REFERENCE NUMBERS: NWU-00178-22-A1

PRINCIPAL INVESTIGATOR: PROF WJH ROESTENBURG

POST GRADUATE STUDENT: IVAN KANES

ADDRESS: 5TH FLOOR, SANDTON GATE, 7 MINERVA AVENUE, GLENADRIENNE, SANDTON, 2196

CONTACT NUMBER: 0714200520

You are being invited to take part in a **research study** that forms part of a Doctoral study. Please take some time to read the information presented here, which will explain the details of this study. Please ask the researcher or person explaining the research to you any questions about any part of this study that you do not fully understand. It is crucial that you are certain of your understanding of the purpose of this research and how you can participate. Your participation is **entirely voluntary** and you are free to refuse to participate. If you refuse, you will not be affected in any way whatsoever. You are also free to withdraw from the study at any point.

This study has been approved by the **Health Research Ethics Committee of the Faculty of Health Sciences of the North-West University (NWU-00178-22-A1)** and will be conducted according to the ethical guidelines and principles of Ethics in Health Research Principles, Processes and Structures and other international ethical

guidelines applicable to this study. It might be necessary for the research ethics committee members or other relevant people to inspect the research records.

What is the purpose of this research study?

This study proposes, using a multi-phased, mixed method approach, the exploration and design of a conceptual A.I. generative chatbot counselling framework to inform an A.I. algorithm, specifically in the field of social work that focuses on the needs of users of a broader employee health and wellness programme.

Why have you been invited to participate?

You have been identified as mental health care practitioner providing online counselling services to mental health care users on regular basis.

What will be requested from you?

Your consent to analyse your online interaction styles, sentiment- and lexicon usage, as well as your counselling approaches- and outcomes will be required.

Will you gain anything from taking part in this research?

There will be no direct gain for your participation in the study. You will indirectly benefit from the development of an intelligent online counselling app.

Are there risks involved in you taking part in this research and what will be done to prevent them?

There are no risks involved in participating in this study.

How will we protect your confidentiality and who will see your findings?

Anonymity is a key feature of this study. All transcripts selected for analysis will be subjected to a process of de-identification.

What will happen with the findings?

All data will be stored in password protected files on the researcher's computer. Data will be stored for 5 years after which it will be destroyed in the prescribed manner.

Future researchers might be able to utilize the data from this study, but they would need the North West University Ethics committee to give them written consent before doing so.

How will you know about the results of this research?

When the study report is finished, you will be notified of the findings, and you will receive an email with a summary of the findings.

Will you be paid to take part in this study and are there any costs for you?

You will not be paid to partake in this stud. There will be no costs incurred upon you, if you partake in this study.

POPIA Disclaimer

We appreciate your right to privacy, therefore we work to guarantee that we adhere to the POPI Act's legal requirements, which govern how we gather, process, store, disclose, and destroy any personal information you have given us.

Is there anything else that you should know or do?

- You can contact Mildred Thomas at 0182992412 or 27381706@nwu.ac.za if you have any additional enquiries.
- You can also contact the Health Research Ethics Committee via Mrs Carolien van Zyl at 018 299 1206 or carolien.vanzyl@nwu.ac.za if you have any concerns that were not answered about the research or if you have complaints about the research.
- You will receive a copy of this information and consent form for your own purposes.

Declaration by participant

By signing below, I ..... agree to take part in the research study titled: Developing an intelligent online counselling app for an employee health and wellness programme.

I declare that:

- I have read this information/it was explained to me by a trusted person in a language with which I am fluent and comfortable.
- The research was clearly explained to me.
- I have had a chance to ask questions to both the person getting the consent from me, as well as the researcher and all my questions have been answered.
- I understand that taking part in this study is **voluntary** and I have not been pressurised to take part.
- I may choose to leave the study at any time and will not be handled in a negative way if I do so.
- I may be asked to leave the study before it has finished, if the researcher feels it is in the best interest, or if I do not follow the study plan, as agreed to.

Signed at (*place*) ..... on (*date*) ..... 20....

....

Signature of participant **Signature of witness**

Declaration by person obtaining consent

I *(name)* ..... declare that:

- I clearly and in detail explained the information in this document to

  .....
- I did/did not use an interpreter.
- I encouraged him/her to ask questions and took adequate time to answer them.
- I am satisfied that he/she adequately understands all aspects of the research, as discussed above
- I gave him/her time to discuss it with others if he/she wished to do so.

Signed at (*place*) ..... on (*date*) ..... 20....

.....
Signature of person obtaining consent

Declaration by researcher

I, Ivan Kanes, declare that:

- I explained the information in this document to ..... **or** I had it explained by ..... who I trained for this purpose.
- I did not use an interpreter
- I encouraged the participant to ask questions and took adequate time to answer them or I was available should the participant want to ask any additional questions.
- The informed consent was obtained by an independent person.
- I am satisfied that the participant adequately understands all aspects of the research, as described above.
- I am satisfied that the participant had time to discuss it with others if the participant wished to do so.

Signed at (*place*) ..... on (*date*) ..... 20....

.....

Signature of researcher

Current details: (23239522) G:\My Drive\9. Research and Postgraduate Education\9.1.5.6 Forms\9.1.5.6_HREC_ICF_Template_Apr2018.docm 25 April 2018

File reference: 9.1.5.6

ANNEXURE 5: INFORMED CONSENT DOCUMENTATION FOR: HEALTH INFORMATICS EXPERT AND MENTAL HEALTH CARE PROFESSIONAL DELPHI PANELISTS

NWU®
NORTH-WEST UNIVERSITY
NOORDWES-UNIVERSITEIT
YUNIBESITI YA BOKONE-BOPHIRIMA

Private Bag X1290, Potchefstroom
South Africa 2520
Tel: +2718 299-1111/2222
Fax: +2718 299-4910
Web: http://www.nwu.ac.za



INFORMED CONSENT DOCUMENTATION FOR: HEALTH INFORMATICS EXPERT AND MENTAL HEALTH CARE PROFESSIONAL DELPHI PANELLISTS

TITLE OF THE RESEARCH STUDY: DEVELOPING AN INTELLIGENT ONLINE COUNSELLING APP FOR AN EMPLOYEE HEALTH AND WELLNESS PROGRAMME

ETHICS REFERENCE NUMBERS: NWU-00178-22-A1

PRINCIPAL INVESTIGATOR: PROF WJH ROESTENBURG

POST GRADUATE STUDENT: IVAN KANES

ADDRESS: 5TH FLOOR, SANDTON GATE, 7 MINERVA AVENUE, GLENADRIENNE, SANDTON, 2196

CONTACT NUMBER: 0714200520

You are being invited to take part in a **research study** that forms part of a Doctoral study. Please take some time to read the information presented here, which will explain the details of this study. Please ask the researcher or person explaining the research to you any questions about any part of this study that you do not fully understand. It is crucial that you are certain of your understanding of the purpose of this research and how you can participate. Your participation is **entirely voluntary** and you are free to refuse to participate. If you refuse, you will not be affected in any way whatsoever. You are also free to withdraw from the study at any point.

This study has been approved by the **Health Research Ethics Committee of the Faculty of Health Sciences of the North-West University (NWU-00178-22-A1)** and will be conducted according to the ethical guidelines and principles of Ethics in Health Research Principles, Processes and Structures and other international ethical guidelines applicable to this study. It might be necessary for the research ethics committee members or other relevant people to inspect the research records.

What is the purpose of this research study?

This study proposes, using a multi-phased, mixed method approach, the exploration and design of a conceptual A.I. generative chatbot counselling framework to inform an A.I. algorithm, specifically in the field of social work that focuses on the needs of users of a broader employee health and wellness programme.

Why have you been invited to participate?

You have been identified as a health informatics expert, or a mental health care professional that has experience in providing mental health care services to mental health care users on a live chat platform on a regular basis and/or you are involved in the development and maintenance of said live chat platform.

What will be expected of you?

You will be requested to participate in four rounds of a Delphi panel. The rounds will consist of and aim to achieve the following:

Round 1: Explore the subject under investigation. Each participant contributes additional pertinent information.

Round 2: Reach consensus on how the group understands the proposed product and its functioning with additional recommendations.

Round 3: Explore disagreements, agreements, and recommendations – identify the underlying reasons and evaluate.

Round 4: Final evaluation, adaptation, and feedback to participants.

You will not be requested to meet but will be requested to raise your opinions independently by e-mail. The Delphi panel will be done to ensure that diversity of opinions is maintained and to limit possible groupthink tendencies.

Will you gain anything from taking part in this research?

There will be no direct gain for your participation in the study. You will indirectly benefit from the development of an intelligent online counselling app.

Are there risks involved in you taking part in this research and what will be done to prevent them?

There are no risks involved in participating in this study.

How will we protect your confidentiality and who will see your findings?

An important component of the Delphi survey is anonymity, which has four major functions: safeguarding the rights of the expert panelists, preventing groupthink, preventing dominance by prominent or influential people, and promoting autonomous decision-making. The Delphi survey is a group communication process whereby the participants may never meet each other or know who has participated in the process. Furthermore, there is the guarantee of anonymity of the participants' individual responses and these are never known to one another. Though this may potentially lead to lack of accountability for the response, this is a possibility not unique to a Delphi study.

Anonymity can be achieved at various levels. Participants may be completely unknown to each other; furthermore, there should be no potential for other panellists to ascribe a response to any individual and the researcher should not reveal their identities. The panelists may be able to identify one other in highly specialized fields, but their judgments and opinions are anonymous, a practice known as quasi-anonymity (Vernon, 2009). The nature of the Delphi survey requires that the researcher follow up the expert panellist, which prevents total anonymity from that aspect as well, rendering it as quasi-anonymous (Keeney, Hasson & McKenna, 2006). Confidentiality is maintained by delinking the data from the person and utilizing codes numbers for returned emails.

Anonymity is a key feature of this study. All communication will be subjected to a process of de-identification.

What will happen with the findings or samples?

All data will be stored in password protected files on the researcher's computer. Data will be stored for 5 years after which it will be destroyed in the prescribed manner.

Future researchers might be able to utilize the data from this study, but they would need the North West University Ethics committee to give them written consent before doing so.

How will you know about the results of this research?

When the study report is finished, you will be notified of the findings, and you will receive an email with a summary of the findings.

Will you be paid to take part in this study and are there any costs for you?

You will not be paid for the study, but you will receive a R400 TakeAlot gift voucher as a token of appreciation for your participation in the study.

Is there anything else that you should know or do?

- You can contact Mildred Thomas at 0182992412 or 27381706@nwu.ac.za if you have any additional enquiries.
- You can also contact the Health Research Ethics Committee via Mrs Carolien van Zyl at 018 299 1206 or carolien.vanzyl@nwu.ac.za if you have any concerns that were not answered about the research or if you have complaints about the research.
- You will receive a copy of this information and consent form for your own purposes.

POPIA Disclaimer

We appreciate your right to privacy, therefore we work to guarantee that we adhere to the POPI Act's legal requirements, which govern how we gather, process, store, disclose, and destroy any personal information you have given us.

Declaration by participant

By signing below, I ..... agree to take part in the research study titled: Developing an intelligent online counselling app for an employee health and wellness programme.

I declare that:

- I have read this information/it was explained to me by a trusted person in a language with which I am fluent and comfortable.
- The research was clearly explained to me.
- I have had a chance to ask questions to both the person getting the consent from me, as well as the researcher and all my questions have been answered.
- I understand that taking part in this study is **voluntary** and I have not been pressurised to take part.
- I may choose to leave the study at any time and will not be handled in a negative way if I do so.
- I may be asked to leave the study before it has finished, if the researcher feels it is in the best interest, or if I do not follow the study plan, as agreed to.

Signed at (*place*) ..... on (*date*) ..... 20....

....

Signature of participant **Signature of witness**

Declaration by person obtaining consent

I *(name)* ..... declare that:

- I clearly and in detail explained the information in this document to

  .....
- I did/did not use an interpreter.
- I encouraged him/her to ask questions and took adequate time to answer them.
- I am satisfied that he/she adequately understands all aspects of the research, as discussed above
- I gave him/her time to discuss it with others if he/she wished to do so.

Signed at (*place*) ..... on (*date*) ..... 20....

.....
Signature of person obtaining consent

Declaration by researcher

I, Ivan Kanes, declare that:

- I explained the information in this document to ..... **or** I had it explained by ..... who I trained for this purpose.
- I did not use an interpreter
- I encouraged the participant to ask questions and took adequate time to answer them or I was available should the participant want to ask any additional questions.
- The informed consent was obtained by an independent person.
- I am satisfied that the participant adequately understands all aspects of the research, as described above.
- I am satisfied that the participant had time to discuss it with others if the participant wished to do so.

Signed at (*place*) ..... on (*date*) ..... 20....

ANNEXURE 6: EDITORIAL LETTER

Marielle Tappan
Wapadrand, Pretoria
Tel 072 474 1158
Email mteditorialinfo@gmail.com

MT Editorial

Date of Edit: 25 October 2024

Edit: Ivan Kanes

DEVELOPING AN INTELLIGENT ONLINE COUNSELLING FRAMEWORK FOR AN EMPLOYEE HEALTH AND WELLNESS PROGRAMME

To whom it may concern,

I, Marielle Tappan, trading under the name MT Editorial, hereby confirm that I am a language editor.

I have extensive experience in the field of language and publishing and received my Bachelor of Information Science in Publishing from the University of Pretoria. I am also a registered member of the Southern African Freelancer's Association.

I hereby declare that the editing done for any client is done with the utmost diligence and the full appreciation of the English language and all of its intricacies, as was done for edited sections of this document.

If there are any other queries, please do not hesitate to contact me.

Kindest Regards,

Marielle Tappan
Owner, MT Editorial
(BIS) Publishing
SAF03058, SAFREA

Marielle Tappan