Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 9, 2025

DEVELOPMENT OF A SECURE AI CHATBOT FOR PERSONAL DATA MANAGEMENT: THE OFFLINE MEMORY MATE ASSISTANT

Tooba Mujahid

MSc Software Engineering, Staffordshire University

missit772@gmail.com

DOI:https://doi.org/10.5281/zenodo.17173416

Keywords
AI Chatbot, Personal Data Management, Data Privacy, Encryption, Natural Language Processing, NLP, Transfer Learning, Offline Features, Question Answering SQuAD 2.0, Flask REST API, Validation Loss, Machine, Secure

Article History
Received: 21 June 2025
Accepted: 03 August 2025
Published: 22 September 2025

Corresponding Author: *
Tooba Mujahid

Abstract
This research presents the development and evaluation of the "Offline Memory Mate Assistant," a secure AI-driven chatbot designed for personal data management. The main goal of this work is to develop a chatbot a secure AI-driven chatbot designed for personal data management, leveraging advanced natural language processing (NLP) specifically transfer learning and encryption techniques to ensure data privacy and protection. Utilizing the Stanford Question Answering Dataset (SQuAD2.0) for training, the research employs the pre-trained deep-set/roberta-base-squad2 model for fine-tuning contextual question-answering capabilities. The encryption method applied is Ferret from Python's cryptography library, providing robust data security through end-to-end encryption. Additionally, the system is implemented with an offline mode to enhance privacy by processing all data locally, eliminating the need for internet transmission. A web-based user interface that is coded in HTML, CSS and JavaScript, Flask REST API for the model's deployment and a Python script for data extraction and encryption. The interaction of the users with the chatbot is quite easy since it incorporates a user interface where the user can select document type and then receives the correct and meaningful response to the queries made. By addressing critical issues such as data privacy, encryption, and regulatory compliance, the study contributes valuable insights into the development of AI-driven systems that enhance user trust and satisfaction while ensuring data security.

INTRODUCTION

In today's digital age, the handling and management of personal data have become critical concerns for individuals and organizations alike. With the increase in the amount of personal information posted and shared online, the protection of this information has become a matter of importance more (Beltrão et al., 2022). It encompasses organizations work performance, social relationships and the routine activities. For storing and to find this info is becoming even more difficult to manage.

This research seeks to address these concerns through the development of an innovative AI- powered solution: which was further known as the Memory Mate Assistant. This is an open AI chatbot whose main function is to allow user's information to be stored and retrieved with a lot of safety, while at the same time, ensuring that the user data is effectively managed across individuals. Memory Mate Assistant signifies the improvement in Personal Information Management and has the vision to offer accurate,

efficient, and easy-to-use tools to organize the user's data.

A major research and knowledge gap with existing solutions is the lack of data reference from the chatbots, especially, when the data is referred entities, along with information about the used document, the data source and the date on which the data was collected (Ghorpade-Aher et al., 2019).

The purpose of this research is to develop the functional prototype of the Memory Mate Assistant and apply state-of-art encryption schemes to protect the users' information. The use of traditional methods is problematic precisely in terms of effectiveness and, practically, the ability to indicate the source and the date of the obtained results.

1.1 Research Problem

Currently existing AI-based personal assistants' application heavily use cloud storage, which is rather insecure, and does not offer the level of protection required for personal data (Sarode et al., 2023). To fill this gap, our Memory Mate Assistant, a secure AI chatbot based on OpenAI's GPT- 4 model, is to be introduced in this research. Besides security the usability, as has been discussed by constitutes the core of the Memory Mate Assistant (Sebastian, 2023a). The goal of this study

is, therefore, to create a simple natural-sensitive user interface where users can enter and store their data and from where they can retrieve this data as well. The layout of the interface shall be made simple and friendly as its core principle will be to be as user-friendly as possible disregarding the experience or know-how of the user. As the officer presents Memory Mate Assistant as a simple and user-friendly tool for organizing personal information, the proposed concept will endow the users with a possibility to navigate through the heavy complexity of its application (Cheng & Jiang, 2020). To the design of the Memory Mate Assistant, this research seeks to facilitate solution design that will enable the user to retrieve data at his/her own convenience. Besides, this feature makes the assistant more reliable and helps to avoid numerous risks connected with data transmission through the Internet.

1.2 Research Aim

To develop a secure AI chatbot, the Memory Mate Assistant, designed for efficient personal data management, ensuring data encryption, personalized interaction, and compliance with data protection regulations.

1.3 Research Objectives

1. To design and implement a secure AI-driven chatbot that effectively manages and retrieves personal data while ensuring data privacy and protection.
2. To integrate advanced natural language processing (NLP) based transfer learning technique for personalized interactions based on the user's habits, preferences, and historical data.
3. To apply encryption methods, specifically Fernet encryption, to secure personal data both during storage and transmission within the chatbot system.
4. To evaluate the effectiveness of offline operation for the chatbot, ensuring enhanced privacy and security by processing all user data locally and avoiding internet transmission.
5. Offer recommendations for future improvements and explore the societal benefits of the chatbot, such as increased data security and improved user experience.

1.4 Motivation

For this reason, the Memory Mate Assistant will utilize the contemporary NLP algorithms. NLP technology helps the chatbot to respond to the user's query or command and make the conversation functional, natural and efficient. Therefore, implementation of efficient NLP algorithms will make Memory Mate Assistant fast and accurate thus enhancing the overall performance of the App. This will make it easier for the users to try to interact with the chatbot in the same way as they would with a real human assistant thus helping in easier and efficient data management (Meske & Bunde, 2023). In the development of the Memory Mate Assistant, there is also a step on testing and evaluation in order to assess the efficiency and reliability of the system as well as its security. Several testing approaches will be used in order to check for possible risks and to assess the effectiveness of the applied encryption methods, ways of interaction, offline mode, and NLP algorithms (Ghorpade-Aher et al., 201)

This study enriches the body of knowledge on AI and PDM by delivering a secure, easy-to-use, and accurate tool for people who want to control their data. The Memory Mate Assistant is a perfect example of how state of the art technologies has been incorporated in the design of an AI chatbot for managing personal information. Through these responses, this work intends to develop a tool that not only has technological features on data security, usability, and offline functionality, yet remains useful to the everyday user (Sreerama & Krishnamurthy, 2022).

1.5 Dissertation Structure

The research work is divided into five major chapters in order to achieve the objectives of the study. The first Chapter: Introduction gives an overall understanding of the topic of the study, the problem statement, the research question, objectives of the study and the importance of the study. It also highlights the area of the research and the questions that the study attempts to answer. Chapter 2: Literature Review presents the current literature, theories, and frameworks on AI chatbots, data privacy and personal data management with gaps that the current study seeks to address. This chapter provides background information and justification of the work, and evaluates the research already conducted. Chapter 3: Research Methods details the methods and approaches used in the study, including the design and development of the Memory Mate Assistant, the techniques employed for data collection, and the processes for data analysis. This chapter also discusses the tools, algorithms, and security measures implemented to achieve the research
objectives. Chapter 4: Design of Artefact presents the findings of the research, analyzing the effectiveness of the developed chatbot in managing personal data securely. It includes a thorough discussion on how the results align with or diverge from the existing literature, addressing the research questions and objectives set out earlier. Chapter 5: Implementation and Testing gives the results of implementation and interface design. Finally, Chapter 6: Conclusion Recommendation summarizes the key findings of the study, highlighting the practical implications, limitations, and recommendations for future research.

Literature Review

2.1 Introduction

This literature review discusses the data privacy issues to other or similar areas of study such as data privacy, protection and security of personal data. Though it focuses on the role of the individual, it reviews ways in which people control their data and looks at methods of data protection. An introduction of model building using Python and some of the most common algorithms in NLP is given. Last, the major issues concerning the creation of personal AI that is secure, private and protected are discussed. This also comprises anonymization and pseudonymization, data storage protection including the methods such as access control, user's authentication, and authorization, policies and approaches of data minimization and getting users' consent, and unstructured data handling.

2.2 Data Privacy Concerns

As the number of applications with artificial intelligence (AI) chatbots increases it means that some questions concerning data privacy issues, data protection, and safety when developing AI are lacking answers AI (Elliott & Soifer, 2022). Ensuring the privacy and security of user data is crucial to building trust and maintaining compliance with relevant regulations. Here are some key considerations:

2.2.1 Data Privacy

Data privacy refers to the protection of personally identifiable information (PII) and ensuring that individuals have control over their data (Sebastian, 2023b). It is crucial to safeguard user privacy when creating an artificial intelligence chatbot. Some important steps include collecting only necessary data, anonymizing information, and getting user permission. The chatbot should only gather data required to function properly for its purpose. Extra details are not required and raise privacy risks. Some information like names must be kept private. The identity of users must remain hidden. Technical methods can help protect identities without losing necessary data. Code encryption or replacing names with ID numbers maintains anonymity (Sebastian, 2023b).

Clear communication with users regarding data collection and use is also essential. People should

understand what information the chatbot needs and how it will be handled. (Song et al., 2023) User agreement must be received prior to gathering any personal details from conversations. Transparency builds trust that privacy is respected. Permissions help ensure users feel in control of their data and know their options.

Security risks such as compromising of data are some of the major issues likely to be realized. Thanks to the massive flow of customers' data, AI chatbots also become appealing to hackers. An attack can grant an unauthorized person access to one's personal information that may be used in defrauding the individual using their information (Majeed & Hwang, 2023). Studies have revealed that as the utilization of the AI systems increases, so does the requirement to protect the information that is processed by the AI systems.

Homomorphic encryption also provides a secure solution for protection of the data since computations can be made on encrypted data without necessarily decrypting them. This makes it possible for AI models to directly work on encrypted information without the need of decrypting the information and thus maintaining the information security of the data set all through the computation process. The outcomes of these computations are then disguised to get the final values, and confidentiality is observed throughout (Gentry, 2009).

In a nutshell, these techniques of privacy-preserving AI are useful to tackle the privacy-related issues of the AI systems. Thus, federated learning, differential privacy, and homomorphic encryption contribute to protecting data privacy and create better security for the AI applications.

2.2.2 Data Protection

Data protection entails undertaking processes for the safety of the user data from outside interference, chance of loss or even being tampered with. Some of these core technologies include; access, access control, storage security, and code retention (Truong et al., 2020). Security measures must be applied on every stage of data processing, and this means that encryption is a mandatory component. Data is enacted during transfer and storage between devices, which makes it possible for only those who possess the right code to comprehend what is in it, although this does not always have to be the intention. Restriction of the access of specific content can help in combating abuse. Using a "need to know" approach to limit in order to decrease the vulnerability to insiders can also be effective. Data that can be contained in user chats, as well as other materials

of a similar nature, can be safeguarded utilizing a backup server and permissions key. This helps minimize on chances of loss or damage as the data can always be restored in case of an occurrence. The concrete time frames to how long it is allowed to retain the information and its deletion once the time is up also safeguarding privacy. It also keeps issues of compliance because it does not retain information which are not essential. These are the possible risks that can be mitigated through the yearly evaluation and the removal of any past data that is no longer needed (Hanupriya and Anto Ramya, 2023). Including aspects about encryption, permissions, storage security, and storage retention makes it possible for a chatbot platform to be developed with an understanding of privacy as it fulfils its role. This balance of design services keeps the users confident of their security.

2.2.3 Security

Security concerns of an AI chatbot can be described as the process by which an AI chatbot is shielded from the various risks that surround it (Meskó and Topol, 2023). Efficient business practices are capable of identifying the defects before they impact on the people. Scheduled reviews aim at looking at the software for exploitable weaknesses. This means that only the permitted user will be able to, for instance, send a message to the bot. The user identification can be optimized by adding the additional steps like one-time password. Secure encoding of the chats keeps the main contents of a conversation concealed during the actual conversation. Response to security threats brings a shield to the users (Meskó and Topol, 2023). It is suggested that an access control strategy should reduce the impact of an incident since it plans the response process.

2.3 Personal Data Management

AI chatbots are computer programs that can converse with the clients using advanced AI tech such as machine learning (ML) and natural language processing (NLP) and simulate human interaction (Daudén-Esmel, Castellà-Roca and Viejo, 2024). It enables the functionality of chatbots for a broad spectrum of tasks ranging from setting up appointments and timers to handling numerous aspects of an individual's life, thereby increasing user efficiency and making life easier and more efficient.

The type of information that was collected by the research article (Jeon, Go and Namgung, 2023) considers the case of Lee-Luda which was an artificial intelligence-Chatbot, that instituted dating counseling service, and which used people's personal information to create and run the chatbot.

But the company which developed Lee-Luda app was punished by the South Korean government for non-compliance with PIPA. The particularities of this case provoke thoughts about the proper application of personal information in the framework of their obtaining for the AI learning data purpose. Based on the facts of the Lee-Luda case, this article considers that there are outstanding influence and constraint effects on South Korea and even the whole world because the using of personal information for other purposes other than the purpose for collection is a universally recognized principle at present. According to the authors, such ethical and legal questions can emerge in other countries as well because there is a conflict between utilizing the collected personal information as training samples for the AI programs and well-established legal frameworks (Jeon Go and Namgung, 2023).

Explorations of this paper is based on the case of AI chatbot; Lee-Luda in South Korea and its contravention of the PIPA. Lee-Luda's construction and administration were based on 10 billion words of personal information obtained at first for a dating counselling service (Jeon, Go and Namgung, 2023). Even though the company intended to use the gathered data to train Machine Learning, the South Korean government fined the company for violating the principles stated in the PIPA. The case of Lee-Luda serves as a precedent in South Korea and has broader implications beyond the country's borders, as the use of personal information for AI learning data challenges the established principle of using personal information solely for its original collection purpose (Jeon et al., 2023).

This paper adopted the legal analysis to discuss the case of Lee-Luda and the Para VI by violating the Personal Information Protection Act (Jeon et al., 2023). Therefore, to determine why the use of personal Information for the learning data in the Lee-Luda program was considered unlawful, the authors discuss the analysis of provisions and regulations as outlined in the PIPA. The study will draw on a wide-ranging analysis of the legal requirements regarding personal information and its utilization in places of learning, the purpose for which the information was gathered, and the possible conflict of interest arising from training of AI from the collected personal data (Jeon et al., 2023).

This point of interest is quite relevant to the given article which largely revolves around the experience of the specific South Korean company Lee-Luda and might not offer a comprehensive insight on the state of personal data security in other countries and among other actors (Jeon et al., 2023). Furthermore, the research does not explain how to create hack-proof chatbots or give clear instructions on the ways personal data protection measures should be incorporated into AI programs.

2.4 Personal data protection Techniques

For the safety of the personal data dealt with by the AI chatbots, several methods have been developed. General security measures that are utilized to protect information include encryption, two factor authentication, and protected storage. Furthermore, it is required to establish the data exchange protection during the communication process: this is why it is essential to use SSL/TLS protocols (Elliott and Soifer, 2022).

Currently, more attention is being paid to privacy-preserving AI methods as the efficient ways to protect data privacy in artificial intelligence (AI) technologies (Majeed and Hwang, 2023). Given that AI is used more and more for dealing with personal data, the problem of confidentiality and data protection becomes vital. In response to these challenges, several techniques have been designed, including Federated Learning, Differential Privacy, and Homomorphic encryption. These methods constituted innovation in

the field of AI, thus aiming at fortifying data security and retaining the performance of AI models (Elliott & Soifer, 2022).

The most frequently used approach to overcome the challenges of protecting personal data is federated learning which enables multiple edge devices to train the model jointly with no raw data transmission. This ensures that data is never shifted from the local devices unlike when using the normal training from a central server. These devices only forward the updates to the central server hence maintaining privacy for each individual but at the same time the model updates from these devices are used to build new models that have been trained by a large number of people (Bonawitz et al., 2019). This is because the data is not easily reachable from other places such as the internet thus making the system safe from hacking and other attacks.

Differential privacy is an important technique that is capable of providing robust privacy mechanism with the help of randomly added noise to the data or even to the results. This noise implies that the inclusion, or exclusion, of any individual data point does not have a significant impact on the outcome of an analysis and thus the contribution of any single data point cannot be discerned and there is no way that certain information can be identified (Abadi et al., 2016).

Differential privacy offers a process that is very valuable in statistical contexts and in situations where data is to be shared while at the same time preserving the privacy of the contributor, but at the same time the data must be aggregated in a way that makes it useful.

2.5 Using Python for Model Building

The research (Sarode et al., 2023) article is concerned with the advancement of real time chatbots with Python and their usage in Python is rich in libraries and frameworks which help developers in creating smart and scalable chatbot applications like NLTK, Flask, and telebot. This paper aims at elucidating on the necessary elements and approaches which can be employed for the construction of a real-time chatbot in Python (Sarode et al., 2023). It explains the methodology of requirement gathering, use case modelling, conversation flow and performance tuning. More specifically, it also discusses the issues of backend services integration, error management, validation, UX design, and the use of NLU for precision and contextual intents recognition. The issues like Security, Privacy and Ethical issues are discussed in the course of developing the chatbot systems. Another factor is information communication technology, which is identified by the authors of the paper as important for the steady state of the chatbot, the paper also identifies prototyping and feedback, and first approach as the factors that should be considered in order to enhance the performance of the chatbot and its perceived use particularly among students (Sarode et al., 2023).

Thus, this study provides a clear demonstration of the opportunities for natural language processing, real-time chatbots created with Python's help, in customer support, automating processes, and offering unique and efficient services to users (Sahoo et al., 2019). The chatbot systems can be designed with the help of Python's libraries and frameworks to understand the intent of the user and answer corresponding to the user's intended message. NLU algorithms and techniques are incorporated in the design of the chatbot to improve the recognition of the users' inputs. The study reveals that, real-time chatbots created with python have the potential of enhancing interactions, performing tasks, and in general enhancing users' satisfaction or experience (Sarode et al., 2023).

However, on the basis of this brief description, it can be assumed that there are some drawbacks and shortcomings in the sphere of the chatbot development process and method. First of all, the article presents the general information concerning the major aspects and the tools needed for creating a real-time chatbot based on Python, but it does not reveal the detailed technical information (Sarode et al., 2023). More studies are required regarding the interactions and information integration of building complex chatbot systems based on more enhanced features such as NLP, Machine learning, and Deep learning models. Moreover, the article does not explore the various restrictions of applying chatbots, especially in real-life scenarios including, how the chatbots understand the inputs of the users, redesigning to accommodate many users, and accommodating for standard and non-standard usages of language (Sarode et al., 2023).

Further studies should quantify the use of the chatbot systems and analyze the problems and possibilities of extensive applicability and reliability in different fields and roles. Finally, the article briefly discusses about the ethical considerations and the explainable AI techniques needed in a chatbot system (Sarode et al., 2023).

A research article (Ghorpade-Aher et al., 2019) reflect that AI chatbots with the help of advance ML & NLP techniques changed the way people maintain their data. These chatbots are built for natural language processing so that the user can get specific services like the booking of an appointment, setting up reminders, and handling several aspects of the personal information. Cognitive technologies such as artificial intelligence and machine learning are beneficial to chatbots as they can process highly voluminous data and offer prompt services to the users thus boosting their efficiency.

2.6 Natural Language Processing (NLP) Algorithms

The implementation of Chatbot based on NLP has grown over the last few years mainly due to the fastest increase in AI and Machine Learning (Nursiyono & Gibran, 2023). NLP or Natural Language Processing is a part of AI that deals with man-machine communication through natural language. The key modules of NLP-based chatbots include the language comprehension, dialog flow control and the language output, although in an interrelated manner as to deliver an easily conversational experience.

Thus, the first step in constructing an NLP based chatbot implies the selection of an adequate NLP model. Existing models like GPT-4 by OpenAI have taken the field to new heights where such models have much higher power for comprehension and generation of text that looks like a human- written text (Sun et al., 2021). The first and often predominant step is the basic pre-training of the model on an extensive text database so that the chatbot can acquire knowledge on language, its

context and semantic structure. It is in this training stage that supervised learning is mostly used as the model is trained with labelled data to discover how it should respond to different inputs (Gupta et al., 2024).

When the model has been trained, the next thing is to incorporate it into a chatbot environment. This entails creating the dialogue management system which refers to the nature of the conversation of the chatbot. The management of the dialogue may be in the form of rule-based or using the machine learning algorithms to predict an appropriate next step given the conversation context (Ghorpade-Aher et al., 2019).

Slightly complex provident applications incorporate reinforcement learning to increment the operations of the application as well as responses with time from emerging interactions.

One of the significant components is to learn how to enable the NLP-based chatbot to recognize the inputs made by the users (Sreerama & Krishnamurthy, 2022). They include the text transformation pre-processing methods including the tokenization of text, lemmatization process, and the distillation of the text until it is in part-of-speech (POS) format as described by Ghorpade- Aher et al. (2019). Also, NER is employed to extract significant information such as names, dates, and locations from the text to help the chatbot learn from the context of the text.

2.7 Challenges in Developing a Secure/Private/Protected Personal AI Assistant

Creating an AI assistant that is less susceptible to leaking users' personal data presents a myriad of issues. These issues are mainly attributed to privacy and security of the consumer data along with the protection of their identities as they seek customized and prompt help (Sebastian, 2023a). Below are the key challenges, identified techniques that pose difficulties, and proposed solutions:

2.7.1 Data Anonymization and Pseudonymization

Ensuring data privacy through anonymization and pseudonymization is essential, but it is complex to model effectively. Some of the suggested steps may pose a challenge to implement since they work with concepts like tokenization and data masking in handling unstructured data. Higher mathematics must be applied to tackle this problem differential privacy and homomorphic encryption are those algorithms. 'Differential privacy" is a method to create noise calculated into data in order to preserve privacy during data analysis (Dong et al., 2022). Homomorphic encryption permits computations on

the data without decrypting the data, hence protecting the
data. Such algorithms have to be embedded into the data processing systems. That is, to keep differentially private while the data is being transformed (Jiang et al., 2023). This helps to conceal personal details of people while at the same time allowing useful data to be derived from the data. Solutions should be applied technically in a careful way that allows both, privacy protection and useful analysis.

2.7.2 Secure Data Storage and Access Controls

Securely storing data while allowing authorized access requires careful planning. Ensuring this is challenging especially when user data is regularly modified. Strong technical solutions need to be incorporated (Tembhare et al., 2019).

Using end-to-end encryption and role-based control (RBAC) can solve these problems. AES-256 encryption protects data at rest with high security. TLS encryption protects the security of data in transit. Integrating RBAC to assign access rights to users based on their roles (Nagrale et al., 2022). This will only allow the affected user to view or edit the profile. Regularly checking access data can help detect unauthorized access attempts. Even if security is compromised, data is still protected. However, with granular permissions, the right people can access the right information. Continuous monitoring further enhances security (Tembhare et al., 2019). A balanced approach that takes into account security protection and user responsibilities can achieve the goals of data protection and authorized use.

2.7.3 User Authentication and Authorization

Verifying identities and authorizing suitable access for users while maintaining usability poses difficulties. Implementing robust solutions to handle these challenges is important (Garcia Valencia et al., 2023). Multi-factor verification (MFA) includes layers of security past passwords but may affect comfort. Utilizing MFA beside OAuth 2.0 authorization addresses this issue successfully. OAuth 2.0 grants third parties constrained get to client information without unveiling qualifications. Counting MFA fortifies protection (Garcia Valencia et al., 2023).

To apply these procedures, MFA ought to be set up requiring extra confirmation steps. OAuth 2.0 permits controlled get to for affirmed applications. Convenience must not be compromised in spite
of the fact that. MFA forms require optimization, so the encounter remains basic for users (Montagna et al., 2023).

Striking an adjust between solid confirmation and authorization versus ease of utilize is key. Receiving specialized measures like MFA and OAuth 2.0 whereas centering on refining interfacing can satisfy both security and involvement objectives reasonably. Clients pick up certainty knowing their accounts are secure whereas holding comfort in verification (Garcia Valencia et al., 2023).

2.7.4 Data Minimization and User Consent

Ensuring data minimization while collecting user data and obtaining explicit user consent can be challenging to balance. It is troublesome to demonstrate frameworks that inquire for as it were vital information and guarantee clear communication with users (Jiang et al., 2023).

Strategically applying the correct arrangements can offer assistance address this adjust effectively. Information minimization calculations help guarantee as it where imperative information is collected (Limna et al., 2023). Assent administration stages skillfully handle energetic client understandings in a straightforward way.

To put such strategies into activity, information capture must be aiming to get fair the basic points of interest required. Assent stages help clients in giving endorsement easily with acumen. Standard assessments of the collection approach confirm consistency with sketched out protection rules and lawful compliance (Cheng & Jiang, 2020).

Achieving focused on information collection matched with well-informed assent requires consolidating specialized helps and diligent appraisals. In the event that executed wisely, such an approach can fulfil person protection interface as well as commerce or investigate objectives through building up well-defined information dealing with conventions (Limna et al., 2023).

2.7.5 Handling Unstructured Data

Managing and securing unstructured data such as text, images, and audio is complex due to its varied nature. Techniques like natural language processing

(NLP) and machine learning (ML) are necessary but can be difficult to implement securely (Sun et al., 2021).

Applying appropriate arrangements warrants thought. State-of-the-art NLP instruments help handle and anonymize unlimited content (Garcia Valencia et al., 2023). For visual or sound records, utilizing secure ML gathering lines counting encryption and anonymization of information in preprocessing helps examination.

execution requires utilizing upgraded NLP like BERT or GPT-3 to translate and anonymize unreservedly shaped content. Pictures and sound require security through ML pipelines including encryption and anonymization some time recently modelling. At long last, affirming security endures from introductory information passage through last show conveyance is key (Ghorpade- Aher et al., 2019).

Deliberate arrangement of calculations custom fitted for the input guarantees strong assessment of random substance. Increasing the forms with scrambled shields jam secrecy. With tirelessness over full improvement, specialized administration of multifaceted fabric ensures clients whereas empowering profitable investigation (Alam, 2012).

2.8 Analysis of Problem/Research Gap

Among these, one of the most important research opportunities is the requirement for a system or chatbot to obtain original data. As far as literature review is concerned, there are theoretical frameworks and guidelines available, still, there is a lack of information about how chatbots can be designed and deployed to help distinguish and locate the primary data sources. This has called for study aimed at identifying methods that could help in formulating systems for processing and managing the genuine user data proficiently.

Moreover, the extent of the study is restricted when it comes to qualitative comparison of a protection of personal data in various countries. The literature is majorly based on the Lee-Luda case in South Korea (Jeon et al., 2023) it does not present all round picture of Personal Data Protection across the globe. More studies are required to filter and discuss various techniques and strategies of personal data protection across the respective countries (Ghorpade-Aher et al., 2019).

First of all, the literature process gives a general description of the conversation development with the help of chatbots using Python but does not contain extended technical details. The given work has revealed some focus areas for research such as the definition of the peculiarities of constructing chatbot systems, for instance use of more complex NLP models, machine learning and deep

learning approaches proposed by Sun et al., (2021). The weakness includes, it does not cover adequately handling of different user inputs, how to accommodate many users, and how to operate within the context of a conversational flow (Barchard & Pace, 2011) Such issues should be addressed in future research for the improvements of the performance and reliability of the chatbot systems.

2.9 Area of Improvement

In order to tackle the understood research gaps in the creation of AI-based chatbots for the management of personal data, this study aims at developing the 'Memory Mate Assistant' for the management and retrieval of personal data while maintaining their confidentiality and security. The solution implies the application of such security measures as encryption concerning the storage of data and the transfer of data. To enhance the perceived utility to the users, one of the modules that shall be incorporated include natural language processing; this allows the chatbot to function optimally based on the user behaviour and activity levels. This approach will still endeavor in ensuring that success in the identification of original data sources for processing shall be afforded to the chatbot Shall address the oversight of general frameworks by providing approaches to the actual handling of data security.

Research Methodology

3.1 Research Design

This study follows a design science research methodology aimed at developing a secure AI chatbot, the "Memory Mate Assistant," for personal data management. The proposed methodology consists in the following steps, which are repeated consecutively until the chatbot fulfils the intended purpose: designing, building, testing, and optimizing. The primary goal is on the design of an efficient and secure AI-based system that can efficiently handle and safeguard users' data.

3.2 Research Method

This research project uses the **Agile method** of project management. It is evident from the definition of Agile that it follows an iterative and flexible approach, which is highly appropriate in this line of research especially when it comes to constructing and developing on the Memory Mate Assistant. This was an Agile process, with its endless communication and modifications according to the supervisor's input and the corrections made throughout the process. This approach also enabled the continuous refinement of the technically and conceptually related aspects of the approach, including security features implemented in the provided chatbot and NLP integration while incorporating new ideas within the process. It is unlike the Waterfall method where it is quite linear and rigid, which makes it rather difficult to carry out corrections and adjustments should there be new discoveries or challenges in the course of implementation at each stage of the process in this project, the Agile framework provided an incremental approach.

The approach used in this study is descriptive and /or analytical, and the data collection method is documentary. This is because they come in handy in identifying emerging trends as well as state- of-the-art technologies, techniques and frameworks in the areas of Artificial Intelligence, Natural Language Processing and Data Protection. This encompasses compilation of data like CVs, habits, likes, and dislikes among others. The objective of these activities is to continuously compile a

profile about each user. Due to the need to quickly search through all the data of the user during interaction, a vector store has been created. In the given vector store emphasis is placed on indexing of the user data so that quick and accurate search and retrieval can be done (Figure 3.1). As for this vector store, there are tools and libraries like FAISS (Facebook AI Similarity Search) that have been employed for its construction.

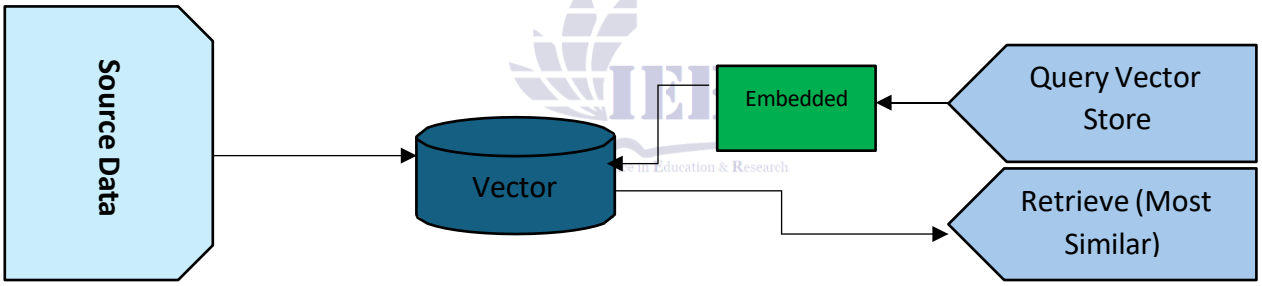

Figure 3. 1 Vector Creation Exemplification
Source: (Self-Created)

3.3 Proposed Research Framework

The process flow given in figure 3.2, presents that this project starts with data collection where the core emphasis is laid on SQuAD 2. 0 question-answer pairs. This preliminary process is important to avoid lack of enough data to feed the chatbot and that is why is commonly referred as foundation data. The process after data has been collected entails pre-processing whereby the data collected is tokenized, undergoes stop word removal and is de-duplicated to improve data quality and usability. Moving forward to model building, after data processing, the next concern is to choose hyper parameters. This entails defining significant attributes which will define the training process including Learning Rate and optimum Model Architecture. What follows hyper parameter selection is encryption and security of the process. Here the use of Fernet module is implemented to get the keys for encrypting and decrypting of the personal data to ensure that the data is secured in the operation of the chatbot. The model then moves in the offline mode, implying that all processing is done offline, which is a measure of increasing security. The model training is at the heart of the workflow, which involves an adaptation of a large language model (LLM) on the processed and secured dataset as well as a conscious and intent-driven learning of the personal data

management requirements. At the end a decision point, after model training, looks at whether validation and testing is required. If validation is needed, further measures are taken to check the efficiency of the model, thus making it more accurate before it is put int use.

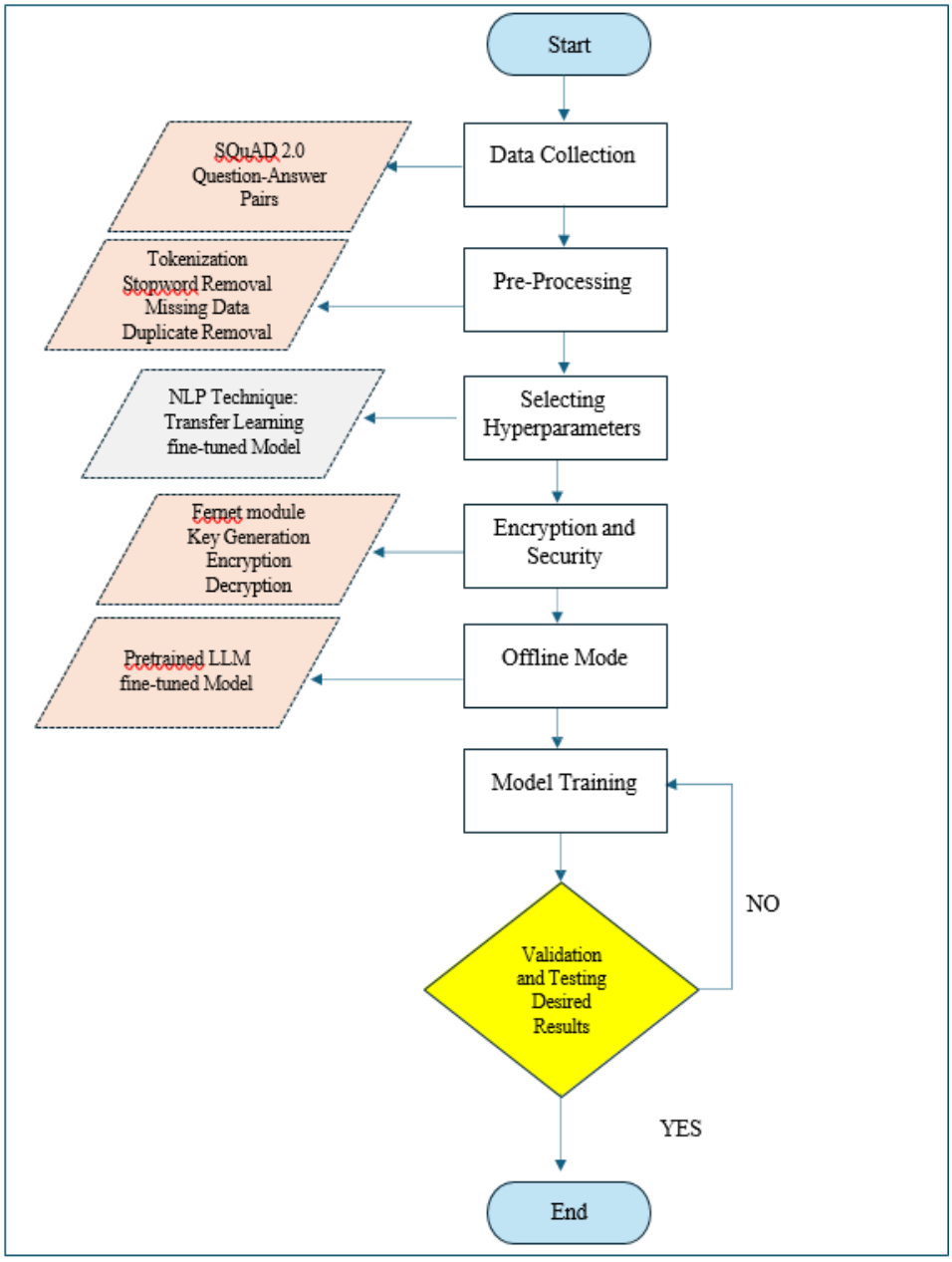

Figure 3. 2 Project Process Flow chart

Design of Artefact

4.1 Data Collection

For this research, the dataset used for training the chatbot was based on **SQuAD 2.0** (Stanford Question Answering Dataset), which contained 130,319 question-answer pairs in the training split and 11,873 question-answer pairs in the validation split. These were critical for fine-tuning the pre-trained

RoBERTa-base model used for question-answering tasks.

The training set of this dataset is made up of 130,319 question-answer pairs and the validation set is made up of 11,873 question-answer pairs. This is because a big data set for training the model is beneficial in the sense that the model can learn numerous patterns and associations between questions and answers in order to (Ali et al., 2023).

4.2 Data Pre-processing

In this research, several data cleaning techniques were applied to ensure the accuracy, consistency, and quality of the data used for training the Memory Mate Assistant chatbot:

1. **Tokenization**: Full text data consist of queries and answers were pre-processed and transformed into individual tokens such as words or parts of the words. This process aids in reducing the volume of text data that needs to be processed hence making it easier for NLP models.
2. **Stopword Removal**: Stems from very frequent words that express very little new meaning in the sentences, such as 'is', 'the' and 'and' were removed in order to eliminate noise in the data set. This makes the model pay more attention to more pertinent terms especially in regard to the management of personal data.
3. **Lowercasing**: The given text data were lowercase transformed to make the analyzing string non case sensitive, i. e. to treat words 'Chatbot' and 'chatbot' identically.
4. **Punctuation Removal**: Amorphous punctuations that do not afford any meaning in a particular sentence were also removed. These codes make the text input easier to interpret and hence, it becomes easier for the model to work with.
5. **Lemmatization**: And the sentences were chopped and converted into words stripping away their prefixes and suffixes. For instance, plural could be modified, such as the word 'running' was changed to 'run'. This makes different forms of a word to be handled similarly; this enhances the generalization of the models.
6. **Handling Missing Data**: If any of the Question answer pairs were having missing values or any entry has been left blank, such Questions and answers were either fitted with suitable values in light of the contents of the same or not included in the data set.
7. **Duplicate Removal**: A process of de-duplication was done to handle redundancy either in question-answer pairs or in the questions themselves so as to eliminate bias.

4.3 Model Building

4.3.1 Selecting the Base Model

In this research, the major method applied in Natural Language Processing (NLP) is transfer learning with pre-trained language models. Specifically, the model utilized is: Specifically, the model utilized is **Deep-set/roberta-base-squad2**. The presented model is a type of a Deep-set model which is a pre-trained model for the question answering tasks that has been fine-tuned on the SQuAD2. 0 dataset, that contains various kinds of question-answer pairs. This model is known to have a good contextual understanding which is very useful for tasks such as question answering (Yan et al., 2022). This is a rather important decision, as this model can be further fine-tuned since it has been pre-trained on a large corpus of different text data.

4.3.2 Setting Hyper parameters

Hyper parameters play a critical role in the training process, influencing how the model learns and how well it performs. The key hyper parameters for this training procedure are as follows:

- **Batch Size (batch size = 96)**: This refers to the number of training examples used in one iteration. A batch size of 96 is chosen to balance computational efficiency and the stability of the model's learning process.

Number of Epochs (n_epochs = 2): The model will be trained for 2 epochs, meaning the entire training dataset will be passed through the model twice. Given the large size of the dataset and the powerful nature of the pre-trained model, two epochs are considered sufficient to fine-tune the model effectively.

- **Base Language Model (base_LM_model = "Roberta-base")**: As mentioned, the Roberta- base model is the backbone of this training process, providing the initial weights and architecture upon which the question-answering capabilities will be fine-tuned.
- **Maximum Sequence Length (max_seq_len = 386)**: This defines the maximum length of the input

sequence, including both the question and the context. Setting this to 386 tokens ensures that most of the input data can be processed without truncation, which could lead to loss of crucial information.

- **Learning Rate (learning rate = 3e-5)**: The learning rate is a critical hyper parameter that controls the size of the steps the model takes during gradient descent. A learning rate of 3e-5 is small enough to allow fine adjustments to the pretrained weights, facilitating effective fine-tuning.
- **Learning Rate Schedule (lr_schedule = LinearWarmup)**: The learning rate schedule used is LinearWarmup, where the learning rate starts from zero and increases linearly during the initial 20% of the training process (warmup_proportion = 0.2), before decaying linearly. This approach helps prevent the model from making large, unstable updates at the beginning of training, allowing for more controlled and effective learning.
- **Document Stride (doc_stride = 128)**: This parameter determines how much the document window moves when it exceeds the maximum sequence length. A stride of 128 tokens ensures that overlapping parts of the text are still considered, improving the model's ability to extract relevant information from the context.
- **Maximum Query Length (max_query_length = 64)**: This limits the length of the input questions to 64 tokens. This restriction helps in managing the input size and ensuring that the model can process the data efficiently.
-

4.4 Data Encryption

4.4.1 Data Encryption Using Python Cryptography (Fernet)

To secure personal data stored and processed by the chatbot, the **Fernet** module from the **Python cryptography library** was used. Fernet is a symmetric encryption method, which means that the same key is used for both encryption and decryption, ensuring that sensitive user data is protected throughout the system's lifecycle.

4.4.2 Key Steps in Data Encryption

Key Generation: Using Fernet we have created a unique encryption key. This key is important during the encryption process of the user's data as well as during the decryption process at some later point in time. The key is kept in a lock box whereby no other person except the usage counter can access it thereby preventing unauthorized use of the elevator.

Encryption: After the key is generated, there is the encryption of user data including names, addresses, emails in the Fernet encryption procedure. This way even if somebody had the misfortune of getting to the data then it would be of no use to him or her since the data is encrypted.

Decryption: In case the chatbot requires the data in order to use it, it decrypts the information using the same Fernet key. This approach makes it possible to restrict the accessible data only for some users or other processes in the application.

4.4.3 The Fernet Encryption Advantages

Fernet encryption has high level of security because it assures the security of data through storage and its transfer which makes it safe for user information. One of its key advantages was the fact that it is fairly easy to implement, because it is a component of Python's open-source library, and did not pose major difficulties to be realized in the system. Furthermore, Fernet also provides end to end encrypted solution where user data is encrypted to maintain privacy and can only decrypted by the right module of the authorized system, so the secrecy is maintained and only individuals who are allowed to have a glance at it can be able to access the information.

4.5 Offline Mode

This Memory Mate Assistant takes advantage of offline operations in order to enhance the confidentiality of user communications and data. Rather than using actual-time online servers, the chatbot is an offline application which downloads a pretrained Large Language Model (LLM) from the Hugging Face library which was fine-tuned for contextual question answering.

Key Steps in Offline Model Implementation

i. Model Download from Hugging Face

Here it is necessary to load the pretrained model either GPT-2 or Roberta-base which can be downloaded from Hugging Face only. This model is pretrained for question-answer type of tasks such as using the SQuAD 2. There was no W0 dataset to

provide context for processing to enable context-based answers.

ii. Local Cache Storage

Once the model is obtained, it is saved in the local storage of the system, thus can be used when the internet connection is not available. This helps in avoiding the potential threats arising from submitting user queries to an online server averting the exposure of the data to such breaches.

iii. Offline Execution

Once the model is cached locally, the chatbot system always loads the model from the cache each time users submits a query. There are no external connections to the online servers; therefore, user interactions take place locally, and the data is safeguarded from attacks from the external networks while it is also processed faster.

iv. Security Benefits of the Offline Mode

The offline mode brings into the table many security features to the table as it eliminates the possibility of user data being intercepted or hacked over the internet. User queries and responses to them are processed locally on the user's device, which means that that personal data leaves the system only if necessary. It provides a high level of privacy as it completely eliminates the use of online servers and third-party systems: all the user data are kept locally.

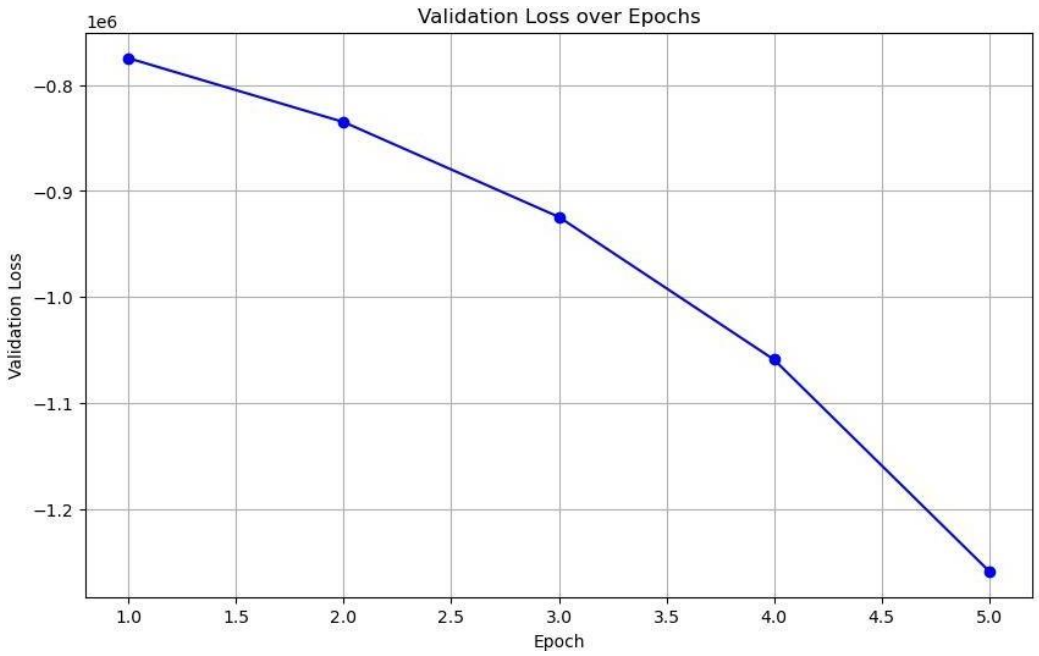

Figure 4. 1 Model Performance and Training Dynamics

Source: (Researcher's Development)

The graph in above figure 4.1 shows the Validation Loss on the y-axis against the number of Epochs on the x-axis to give an understanding of the model's performance and training progress. Validation loss is one of the most important metrics that indicate how well the model can perform on unseen data and is calculated on a validation set that is not used during training to prevent overfitting. In the beginning, the validation loss is around -0. 8. This is the initial value that is set, and it provides a starting point of the model's performance before any changes are made through the learning process. Over time, the validation loss drops steadily as depicted in the graph. This is a negative trend which shows that the model is slowly but surely getting better and is able to generalize better on the validation set. This is due to the fact that the curve has a relatively smooth bend, which

indicates that the validation loss is gradually decreasing and approaching a certain lower bound, thus meaning that the model is gradually correcting the errors it makes on the validation set.

GitHub link for Code:
Here's the link of GitHub for Memory Mate offline assistant artefact code: https://github.com/Tooba2024/memory-mate-assistant

Implementation and Testing

5.1 Model Training Results

The configuration of the Roberta-base model and its training and fine-tuning on the SQuAD 2. Based on the 0 dataset, the model showed its efficiency in answering complex questions, which proves the effectiveness of the model (Liu et al., 2024). The training process depicted that the validation loss was gradually decreasing across the epochs which show that the model is able to learn from the training data and able to perform well on the new instances. This trend implies that the selected hyper parameters for instance, the batch size, learning rate and maximum sequence length were well fine-tuned to enhance the model without overcompensating.

One of the most important findings was the model's capacity to determine when it is not possible to provide an answer, which is a very specific issue of the SQuAD 2. 0 dataset. This aspect of the model is very important since it shows that the model is able to understand context and is aware of the fact that there are some ambiguities or lacks in the given data (Lilian et al., 2021). The model was also able to work with a large number of questions and different contexts from simple questions that have yes or no answers to the more complex ones that required more than simply reading the text.

The second split called validation split consisted of 11,873 question-answer pairs which were used to assess the effectiveness of the model for practical purposes. This particular model's performance on this validation set was good, given that it was able to give correct answers most of the time and even identify questions that it cannot answer. This result further supports the idea of the fine-tuning process and shows the versatility of the roberta-base model in particular. The detailed summary of the results is given in table 5.1 below

Table 5. 1 Summary of the Results

Metric	Value	Description
Model Name	deepset/roberta-base-squad2	Pretrained model used for question-answering
Dataset	SQuAD 2.0	Stanford Question Answering Dataset
Batch Size	96	Number of samples processed before model update
Number of Epochs	2	Number of complete passes through the training dataset
Base Language Model	roberta-base	The underlying language model for fine-tuning
Max Sequence Length	386	Maximum length of input sequences
Learning Rate	3e-5	Rate at which the model updates weights
Learning Rate Schedule	LinearWarmup	Type of learning rate schedule used
Warmup Proportion	0.2	Proportion of warmup steps before learning rate stabilization
Document Stride	128	The stride length for splitting long documents
Max Query Length	64	Maximum length of queries
Training Split Size	130,319 question-answer pairs	Size of the training dataset

Validation Split Size	11,873 question-answer pairs	Size of the validation dataset
Validation Loss	Decreasing with epochs	Indicates model generalization over unseen data

5.2 Critical Evaluation

5.2.1 Generalization Capability

This was because the validation loss of the model kept on decreasing with increases in epoch, hence implying the model has a good ability to generalize. This can be attributed to the generalization capability of the model which is the ability to find the general patterns from the training set then apply it on the new set of data which is vital in real life applications. This is a crucial characteristic for the AI systems that are employed in various contexts since they respond to many users' queries that the systems could not encounter during training. This capability keeps the chatbot more reliable and robust as it is in a position to answer a new and/ or unfamiliar question correctly.

5.2.2 Handling Unanswerable Questions

It is also worth noting that one can find out that the model is capable of approximating that a question has no right answer. This is important specially to regain the user's trust and satisfaction in the AI systems. Therefore, it makes users avoid getting misleading or irrelevant information by identifying cases when a query is ambiguous or incomplete and by reacting to it accordingly. This also makes the system seem more usable by controlling the expectations of the users and at the same time it makes the system seem more believable. In doing so it measures the fidelity of the model to respond appropriately to rich and detailed conversational structures.

5.2.3 Hyperparameter Optimization

However, it is prudent to point out that, the process of optimizing the hyperparameters was very sensitive to improving the performance of the model. Experiments were carried out with changes in the learning rate schedule and the batch size and also the maximum length of sequences so as to have an effective learning without necessarily over-fitting. This process of optimization does play the important role to ensure that the model was in a good standing for then doing the training and validation dataset. Such selection of the parameters was stable and performant, proving how important hyper parameters tuning is while building the efficient machine learning models.

5.2.4 Scalability and Practical Application

Training of the model was done on more than 130,000 question-answer pairs and validated on nearly 12,000 pairs, which shows that it is extendable. The findings suggest that it is possible to apply the approach to other datasets and areas, which proves the versatility of the NLP tasks. This ability proves useful in extending the scope of the model to other scenarios and needs, and therefore the model can be considered as useful for a number of applications beyond the scope of this study. The positive results achieved by applying transfer learning using the roberta-base model show the possibility of its further application and can be considered as a precedent for further research in the field of AI-based question-answering systems.

5.3 Testing

In the implementation and testing phase, the chatbot is tested to the extent of verifying the ability of the system in retrieving information from indexed documents. The system starts with user query where the retrieval process is initiated. The documents which are to be indexed may be in Word or PDF format and they are stored in the document indexes where the retriever will search for the relevant content, highlighted in figure 5.1. In the testing phase we measure the ability of the retriever to accurately find the right context given a range of user queries. After acquiring the context, the system sorts it depending on its relevance, thus generating a ranked context. The generator, backed by a large language model (LLM), forms this ranked context into a meaningful answer. During the whole testing process, we check how the system is capable of

providing correct and proper answers. It is also important to mention that response time, context retrieval accuracy, and generated answer relevance is evaluated regularly to guarantee the best performance before the final launch.

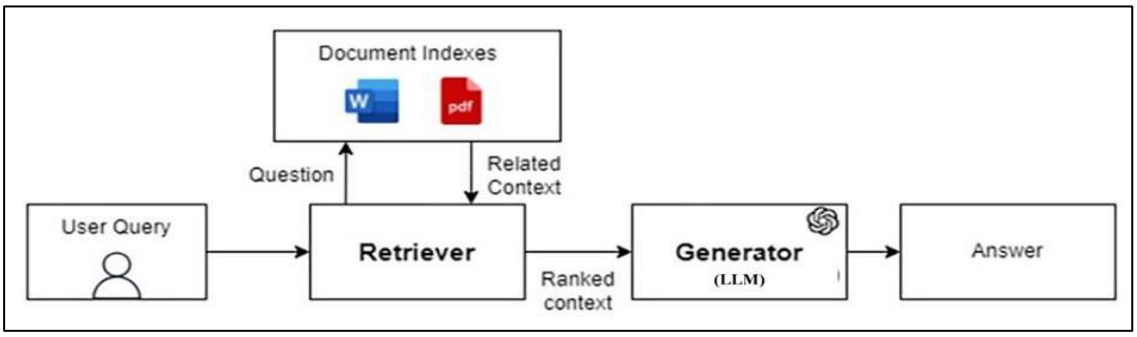

Figure 5. 1 Implemented framework for offline memory mate development (Source: Self-Created)

5.4 Interface Design

The Memory Mate Assistant web-based application has been designed with the aim of the user interface and ease of use. To create the interface design, the system uses HTML, CSS and JavaScript for the front-end making it easy to use and friendly to the end users. The Flask REST API is used for the integration of the proposed AI model, and it offers an easy means of deploying the model as well as facilitating the communication between the user interface and the back-end system. For data extraction and encryption, a simple Python script is used so that everyone's data is protected during the whole process.

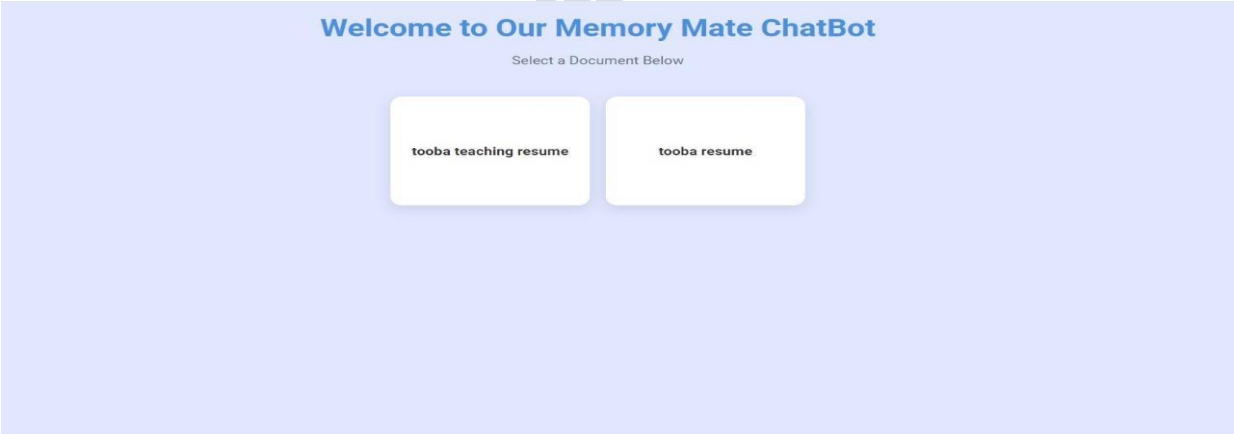

Figure 5. 2 Interface layout of Memory Mate Assistance
(Source: Researcher's Development)

The introduction of the interface sees a welcome message at the top of the page for the users to engage with the chatbot. Strictly after the greeting, there are two buttons with labels 'tooba teaching resume' and 'tooba resume' as given in above figure 5.2. These buttons assist the users in identifying the document they need more information about the design is quite simple with the use of light blue background which makes the text easily readable and the application appealing. This color scheme also gives a sense of comfort and friendliness as the user thus making the chatbot more friendly to be around.

The first and the most important part of the interface contains the conversation between the user and the chatbot as a chat. User queries are placed in light green boxes while the chatbot's responses to these queries are directly written below the boxes in green color which helps in creating a conversation like format (Figure 5.3). The questions that are often asked include Contact information for instance emails,

phone numbers, and work experience. This makes the chatbot responsive and give brief answers to the users' questions to avoid prolonging their search for information.

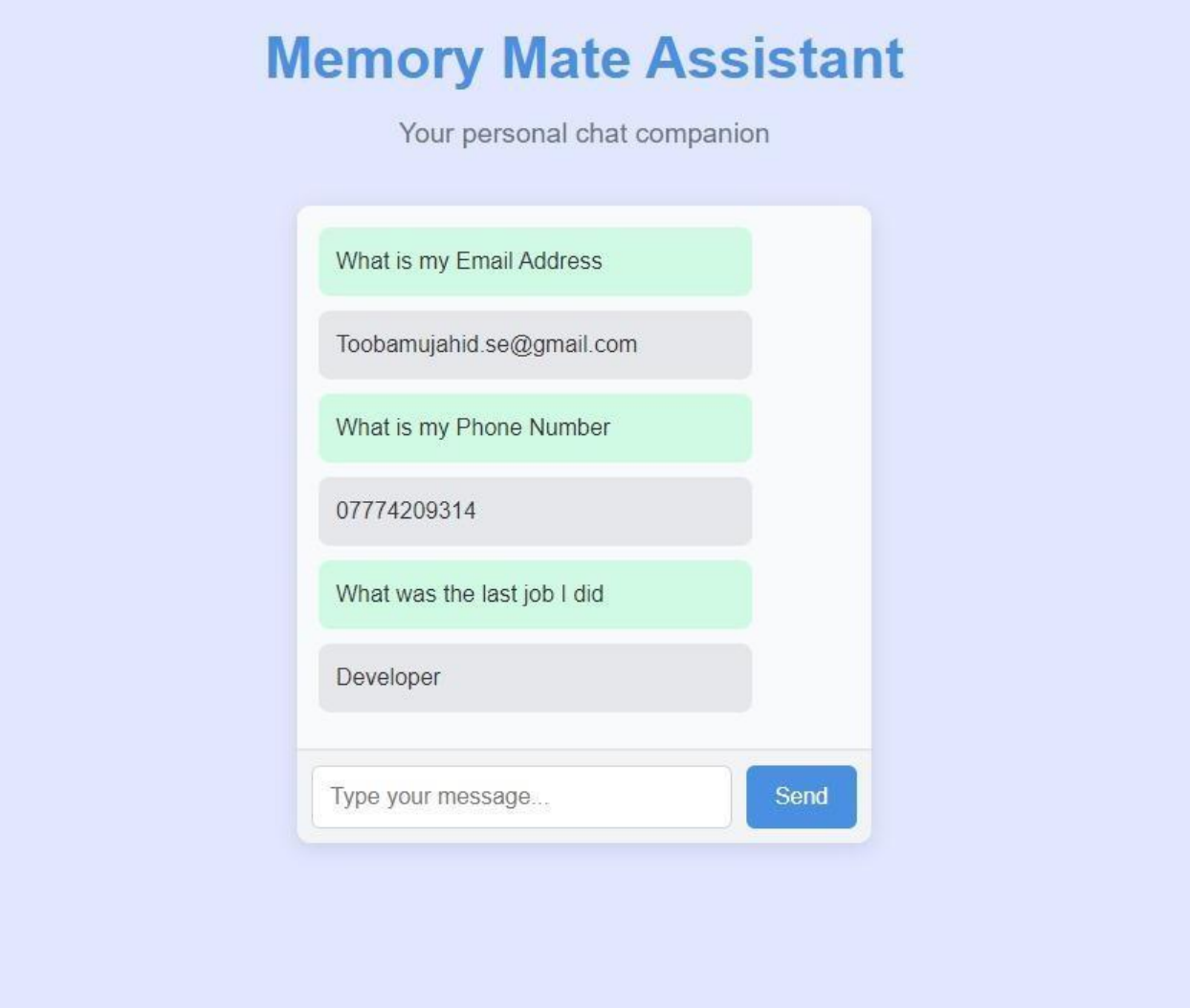

Figure 5. 3 Response Interface, Testing the Artefact

(Source: Researcher's Development)

At the bottom of the interface there is an empty textbox with the label "Type your message." where the new queries can be typed (Figure 5.3). This is followed by a "Send" button which makes it easy for the users to post their questions. The use of the input field alongside the chat feature allows the users to interact with the chatbot real time thus creating a form of engagement that is personal and interactive.

In particular, Flask REST API helps in the interaction between the web-based interface and the model to make the chatbot work effectively and respond to the queries introduced to it. The Python script which is used for the purpose of data extraction and encryption is quite vital in the sense of data protection of the users. The process of encryption is used for any personal data that can be shared in the course of interactions and thus protect it from being intercepted by unauthorized individuals.

To that effect, the Memory Mate Assistant takes into consideration the usability and security aspects in the

design of its interface thus making it suitable for personal data management. With a clear and simplified interface and the proper management of users' information, the Memory Mate Assistant achieves its purpose as an effective AI-based chatbot.

Conclusion and Recommendations

6.1 Overall Conclusion

The purpose of this study was to design a secure AI-based chatbot for effective data management, which is the Memory Mate Assistant. Due to the application of state-of-art NLP methods and strict measures of data security and privacy, the chatbot was developed to offer the personalized communication with the user according to his history and preferences. For question answering, the GPT2 for Question Answering was used and the RoBERTa-base architecture was used for the system which was then fine-tuned on SQuAD 2. The proposed approach utilizes a dense 0 dataset to improve the model's capability of understanding and answering user's query.

One of the critical advancements in this project is the implementation of strong encryption techniques, specifically utilizing Python's cryptography library with Fernet encryption. It guarantees the privacy of all user information throughout the storage process and the transmission process as well. This kind of encryption ensures user confidentiality is maintained by making sure that data can only be decrypted by genuine systems therefore cannot be accessed by other unauthorized persons.

Also the inclusion of the offline mode greatly improves security as no data transmission through the internet is required for the data to stay in the user's local system. On top of that, by processing all client queries and data locally it also ensures that there are no instances where information is intercepted and or leaked whilst being transmitted. All these features of encrypting and offline mode not only guarantee the high security of users' data but also contribute to the increase in popularity and confidence in artificial intelligence systems. That the above measures have been implemented successfully serves to support the need to adopt privacy-minded solutions in the creation of AI systems meant for handling personal data.

This research showed that the chatbot provided a good level of user engagement while at the same time ensuring the data security. This was observed through validation loss metrics wherein the model's generalization capacity was enhanced across epochs meaning that the model training and optimization was effective. On the whole, this work effectively responds to the difficulties of building secure and easy-to-use AI chatbot for the management of personal data, as well as provides valuable guidelines for future developments in the field of data protection, regulation, and safe AI application.

6.2 Societal Benefits and Contributions

The Memory Mate Assistant chatbot helps in improving the society in several ways and plays a vital role in the management and protection of personal data. For those who are using the system, especially the individuals it is easy and safe to store and retrieve information in the chatbot. For instance, the professionals can employ the system to retrieve and modify information from their resumes or CVs while at the same time protecting their information. It also contributes to the protection of privacy, as well as increase the productivity and decrease the likelihood of information leakage.

Encryption also gives a guarantee that personal information is safe all the time through storage and even in the process of transmitting it. This is especially important in the current world where cases of data theft and identity theft are very rampant. For instance, think of an application in the healthcare domain, where a user might want to know more about his or her record stored in the system. With encryption, even if the intruder has gained the access to the data, it will not be intelligible in a readable form to the user, thus the privacy of user.

By keeping all data encrypted, this system helps in the fight against cybercrimes such as hacking and phishing, which are growing threats to individuals and organizations. Encryption ensures everyone has a very hard time trying to infiltrate and use or misuse the data that is transmitted between people.

Also, the chatbot has a simple interface that enhances access and inclusion of users. It has elements such as buttons and interfaces which are easy to locate and use by any user including those with disabilities or those with low technical knowledge to use digital services. For instance, the elderly user with no or little technical knowledge can use the chatbot to manipulate their data or get information from the digital documents without necessarily having to seek for help from the technical team. This increases the users' involvement and guarantee that a large number of people can use the digital platforms.

The developers of the Memory Mate Assistant have also ensured that the application adheres to the international data protection standards hence making the AI-based tool a perfect reference. Thus, obeying the rules of the international data protection, the chatbot offers a model that can be used by other systems and areas. This is especially useful for policy makers and human rights activists who are in the process of drafting or fine-tuning data protection laws. The chatbot's implementation can be considered as an example of the proper approach to data protection in the digital environment.

Offline data processing is advantageous given that it does not expose important information through transmission in problematic communicating networks. During the period within which the system is compromised or a data breach is achieved, offline mode restricts the attackers from accessing live data. For instance, in cases of a cyber-security attack, offline data storage implies that the perpetrators cannot get hold of real-time personal information and thus, containing harm.

6.3 Recommendations for Future Work

Having gone through the process of developing the memory Mate Assistant, there are some challenges that I have observed which may be seen as areas that future researchers can build on. This project has helped me to gain an understanding of the issues related to the creation of a secure AI-based chatbot for managing a person's data. The following recommendations are made to help the future researchers who wish to further develop this technology or work on similar projects.

6.3.1 Scale and performance improvement

Another problem that I faced was how to manage the chatbot's effectiveness when the number of queries increased significantly. When the users of the system increase, the amount of data that has to be interacted with and managed by the chatbot becomes much larger. It is my suggestion that future studies should concentrate on enhancing the system's design with a view of accommodating huge volumes of transactions without affecting the system's response time or data retrieval efficiency. Other approaches including distributed computing, parallel computational paradigms and enhanced caching solutions may be employed in order to enhance scalability.

6.1.1 Risks and Ethical Concerns in Bias of Artificial Intelligence

Thus, during the project, I understood the necessity of ethical concerns, especially in avoiding prejudice in the AI decisions. Although measures were taken to reduce bias, there are issues as to how best to make the chatbot's response polite, non-biased and appropriate for all users. First, I would like the future researchers to focus on the implementation of ethical AI and perform reviews of the AI model's behavior. Some of the current measures include fairness-aware machine learning, and the creating of the explainable artificial intelligence.

6.1.2 Recommendation on Encryption

Recommendations for Further Research On the basis of this study, further research should focus on; Improved encryption techniques other than Fernet to improve on the security of personal data. For example, the employment of Elliptic Curve Cryptography (ECC) or Homomorphic Encryption could serve more security and efficiency.

6.1.3 Recommendation on Offline Feature

Introducing asynchronous classroom and/or office workstation and computer-tailored home workstation Future improvements should incorporate a new offline-online type of class/office workstation and home workstation tailored to computer communication. This hybrid model could make it possible for the Memory

Mate Assistant to remain offline at least to perform basic operations while at the same time exhibiting high privacy and data security; the user can always go for online mode to obtain additional features or updates as a result of the online capacity.

6.1.4 Improve User Interface and Experience

Although I was able to design a good user interface, I faced some difficulties as to how to make the design fully accessible and understandable to the user regardless of disability. Further studies must be directed towards improving the UI and UX and such measures as usability testing and user feedback should be employed. Some of the ways that can make a big difference in the user experience include adding adaptive UI elements that change as per the user's behavior and preferences.

6.3.2 Handling Multimodal Inputs

The chatbot was mainly programmed to entertain text inputs and feedback which made it difficult to understand and respond to queries that include images, voice or any other form of data. This limitation made it less effective, especially in the situations where the user interaction is more complex.

Further research should consider the implementation of multimodal learning models which are capable of analyzing multiple inputs at the same time. This could include integrating NLP with CV and speech recognition to come up with an integrated interaction system. The utilization of transformer models, such as CLIP (Contrastive Language-Image Pre-training) that is capable of working with text and image data may be useful.

REFERENCES

Alam, M. (2012).
Cloud Algebra for Handling Unstructured Data in Cloud Database Management System. *International Journal on Cloud Computing: Services and Architecture*, *2*(6). https://doi.org/10.5121/ijccsa.2012.2603

Ali, Y. A., Awwad, E. M., Al-Razgan, M., & Maarouf, A. (2023).
Hyperparameter Search for Machine Learning Algorithms for Optimizing the Computational Complexity. *Processes*, *11*(2). https://doi.org/10.3390/pr11020349

Amisse, C., Jijón-Palma, M. E., & Silva Centeno, J. A. (2021).
Fine-tuning deep learning models for pedestrian detection. *Boletim de Ciencias Geodesicas*, *27*(2). https://doi.org/10.1590/S1982-21702021000200013

Barchard, K. A., & Pace, L. A. (2011).
Preventing human error: The impact of data entry methods on data accuracy and statistical results. *Computers in Human Behavior*, *27*(5). https://doi.org/10.1016/j.chb.2011.04.004

Beltrão, G., Paramonova, I., & Sousa, S. (2022).
User Interface Design for AI-Based Clinical Decision-Support System. *Iberian Conference on Information Systems and Technologies, CISTI*, *2022-June*. https://doi.org/10.23919/CISTI54924.2022.9820378

Cheng, Y., & Jiang, H. (2020).
How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use. *Journal of Broadcasting and Electronic Media*, *64*(4). https://doi.org/10.1080/08838151.2020.1834296

Dong, J., Roth, A., & Su, W. J. (2022).
Gaussian differential privacy. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *84*(1). https://doi.org/10.1111/rssb.12454

Elliott, D., & Soifer, E. (2022).
AI Technologies, Privacy, and Security. *Frontiers in Artificial Intelligence*, *5*. https://doi.org/10.3389/frai.2022.826737

Garcia Valencia, O. A., Suppadungsuk, S., Thongprayoon, C., Miao, J., Tangpanithandee, S., Craici, I. M., & Cheungpasitporn, W. (2023). Ethical Implications of Chatbot Utilization in Nephrology. In *Journal of Personalized Medicine* (Vol. 13, Issue 9). https://doi.org/10.3390/jpm13091363

Ghorpade-Aher, J., Kontamwar, R., Kukreja, S., Karpe, T., & Kakkad, S. (2019). An overview of NLP based Chatbots. *Universal Review*, *VIII*(II).

Jeon, S. J., Go, M. S., & Namgung, J. H. (2023). Use of personal information for artificial intelligence learning data under the Personal Information Protection Act: the case of Lee- Luda, an artificial-intelligence chatbot in South Korea. *Asia Pacific Law Review*, *31*(1). https://doi.org/10.1080/10192557.2022.2117483

Jiang, H., Pei, J., Yu, D., Yu, J., Gong, B., & Cheng, X. (2023). Applications of Differential Privacy in Social Network Analysis: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, *35*(1). https://doi.org/10.1109/TKDE.2021.3073062

Lilian, J. F., Sundarakantham, K., & Shalinie, S. M. (2021). QeCSO: Design of hybrid Cuckoo Search based Query expansion model for efficient information retrieval. *Sadhana - Academy Proceedings in Engineering Sciences*, *46*(3). https://doi.org/10.1007/s12046-021-01706-0

Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching*, *6*(1). https://doi.org/10.37074/jalt.2023.6.1.32

Liu, J., Lin, S. X., & Wang, S. P. (2024). Question- Answer Pairs Generation Based on Interaction- Guided Joint Abstractive Model. *Jisuanji Xuebao/Chinese Journal of Computers*, *47*(2). https://doi.org/10.11897/SP.J.1016.2024.00251

Meske, C., & Bunde, E. (2023). Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection. *Information Systems Frontiers*, *25*(2). https://doi.org/10.1007/s10796-021-10234-5

Montagna, S., Ferretti, S., Klopfenstein, L. C., Florio, A., & Pengo, M. F. (2023). Data Decentralization of LLM-Based Chatbot Systems in Chronic Disease Self-Management. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3582515.3609536

Nagrale, V., Yalij, M., & Kumar, A. (2022). Data Security of Dynamic and Robust Role Based Access Control from Multiple Authorities in Cloud Environment. *International Journal for Research in Applied Science and Engineering Technology*, *10*(5). https://doi.org/10.22214/ijraset.2022.42489

Özçift, A., Akarsu, K., Yumuk, F., & Söylemez, C. (2021). Advancing natural language processing (NLP) applications of morphologically rich languages with bidirectional encoder representations from transformers (BERT): an empirical case study for Turkish. *Automatika*, *62*(2). https://doi.org/10.1080/00051144.2021.1922150

Rajpurkar, P., Zhang, J., Lopyrev, K., & Liang, P. (2016).
SQuAD: 100,000+ questions for machine comprehension of text. *EMNLP 2016 - Conference on Empirical Methods in Natural Language Processing, Proceedings*. https://doi.org/10.18653/v1/d16-1264

Sarode, Prof. V., Joshi, B., Savakare, T., & Warule, H. (2023).
A Real Time Chatbot Using Python. *International Journal for Research in Applied Science and Engineering Technology*, *11*(5). https://doi.org/10.22214/ijraset.2023.53453

Sebastian, G. (2023a).
Privacy and Data Protection in ChatGPT and Other AI Chatbots. *International Journal of Security and Privacy in Pervasive Computing*, *15*(1). https://doi.org/10.4018/ijsppc.325475

Sebastian, G. (2023b).
Privacy and Data Protection in ChatGPT and Other AI Chatbots: Strategies for Securing User Information. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4454761

Song, M., Zhu, Y., Xing, X., & Du, J. (2023).
The double-edged sword effect of chatbot anthropomorphism on customer acceptance intention: the mediating roles of perceived competence and privacy concerns. *Behaviour and Information Technology*. https://doi.org/10.1080/0144929X.2023.2285943

Sreerama, J., & Krishnamoorthy, G. (2022).
Ethical Considerations in AI Addressing Bias and Fairness in Machine Learning Models. *Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (Online)*, *1*(1). https://doi.org/10.60087/jklst.vol1.n1.p138

Sun, D., Zhang, X., Choo, K. K. R., Hu, L., & Wang, F. (2021).
NLP-based digital forensic investigation platform for online communications. *Computers and Security*, *104*. https://doi.org/10.1016/j.cose.2021.102210

Tembhare, A., Sibi Chakkaravarthy, S., Sangeetha, D., Vaidehi, V., & Venkata Rathnam, M. (2019).
Role-based policy to maintain privacy of patient health records in cloud. *Journal of Supercomputing*, *75*(9). https://doi.org/10.1007/s11227-019-02887-6

Toledo, G., & Marcacini, R. (2022).
Transfer Learning with Joint Fine-Tuning for Multimodal Sentiment Analysis. https://doi.org/10.52591/lxai202207173

Yamanaka, S., Nagatomo, T., Hiraki, T., Ishizuka, H., & Miki, N. (2022).
Machine-Learning- Based Fine Tuning of Input Signals for Mechano-Tactile Display. *Sensors*, *22*(14). https://doi.org/10.3390/s22145299

Yan, A., McAuley, J., Lu, X., Du, J., Chang, E. Y., Gentili, A., & Hsu, C. N. (2022).
RadBERT: Adapting Transformer-based Language Models to Radiology. *Radiology: Artificial Intelligence*, *4*(4). https://doi.org/10.1148/ryai.210258