AI-powered bilingual conversational agent for e-learning

Authors

Masab Saleem Umar
Wasiq Nabi Baksh
Hammad Alam

Supervisor: Mr. Ijaz Ullah

Final Year Project Report submitted in partial fulfillment of the requirements for the Degree of BSDS (Hons.)

INSTITUTE OF MANAGEMENT SCIENCES, PESHAWAR
PAKISTAN

Session: 2021-2025

Certificate of Approval

I, certify that I have read the report titled: **AI-powered bilingual conversational agent for e-learning**, by **Masab Saleem Umar, Wasiq Nabi Baksh and Hammad Alam**, and in my opinion, this work meets the criteria for approving the report submitted in partial fulfillment of the requirements for BSDS (Hons.) at Institute of Management Sciences, Peshawar.

Supervisor: Mr. Ijaz Ullah

Lecturer

Signature: _____

Coordinator BSDS (Hons.): Dr. Bahar Ali

Assistant Professor

Signature: _____

Coordinator FYP: Mr. Omar Bin Samin

Lecturer

Signature: _____

Declaration

We, **Masab Saleem Umar, Wasiq Nabi Baksh and Hammad Alam**, hereby declare that the Final Year Project Report titled: **AI-powered bilingual conversational agent for e-learning** submitted to FYP Coordinator and R&DD by us is our own original work. We are aware of the fact that in case, our work is found to be plagiarized or not genuine, FYP Coordinator and R&DD has the full authority to cancel our Final Year Project and We will be liable to penal action.

Masab Saleem Umar

BSDS (Hons.)

Session: 2021-2025

Wasiq Nabi Baksh

BSDS (Hons.)

Session: 2021-2025

Hammad Alam

BSDS (Hons.)

Session: 2021-2025

Dedication

We dedicate this Final Year Project to our parents and teachers, who have always supported and helped us in every aspect of life.

Acknowledgement

All the praise to Allah that induced the man with intelligence, knowledge, and wisdom. Peace and blessing of Allah be upon the Holy Prophet who exhort his followers to seek for knowledge from cradle to grave.

Foremost, We would like to express our sincere gratitude to our supervisor Mr. Ijaz Ullah for his continuous support, patience, motivation, enthusiasm, and immense knowledge. His guidance helped us throughout the project. Last, but not the least, We would like to thank our parents for supporting us morally and spiritually throughout our life.

Abstract

Your abstract comes here. (Required Word Count: 200 to 300 Words).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A chatbot is a program that mimics human conversation. It constantly works taking and answering requests. How well a chatbot can understand a human question and provide the right answer determines how effective it is. The idea of chatbot's started with early computer science, especially Alan Turing's work in the 1950s. He tested if a user could tell the difference between human and machine talk, the goal was to create smart systems that could act like humans [1]. Although chatbot tech has improved a lot, it still struggles to chat as fluently and about as many topics as expected [2]. Scientists are testing different ways to improve Chabot's. The chatbot story began in 1966 with ELIZA, a simple program that fooled people into thinking it was human [3]. After ELIZA, chatbot tech grew, adding features like voice recognition and emotion recognition. In 2001, Smart-Child made a big splash, working on popular systems like MSN and AOL [4]. In 2006, IBM built a Chabot to compete on TV show "Jeopardy," using NLP to find information quickly, but it struggled with long chats [5]. Around 2014, virtual assistants like Cortana, Siri, and Alexa revolutionized chatbot's by allowing back-and-forth chat and logical thinking [6]. At the same time, Facebook's Messenger let businesses make chatbot's without needing AI, making chatbot tech easier to access. Now, more and more businesses are using chatbot's to better connect with customers [7].

1.1 Overview

Technology has advanced quickly and changed many areas, like education. But, in countries such as Pakistan, it's still difficult to have access to basic education. Here, lots of people mainly speak Urdu and battle economic problems and issues connecting to the internet. As the Pakistan Education Statistics 2021-22 report shows, Pakistani education isn't equal, especially between city and countryside areas. For example, about 26.2 million kids who are old enough for school aren't attending school, many of them from areas that lack development [8].

In today's world, being connected to the internet is vital, especially in educa-

tion. This need became even clearer during the COVID-19 pandemic when schools started using online learning. However, a steady internet connection isn't always accessible in rural areas. Dawn reported in 2020 that a Senate committee talked about this issue. They said students from far-off places often have trouble joining online classes. This problem comes from unstable internet connections and the high price of tech gadgets. The government stresses on a complete plan for tackling the problems, pointed out by the Senate's IT committee. Proposals consisted of cheaper internet plans, cost-effective tech devices, and setting up data centers in less-served areas to boost online learning. Due to school interruptions during the pandemic, the Higher Education Commission (HEC) started projects to enhance digital reach for learners. The collaboration of HEC with telecom firms resulted in student-specific internet packages, yet hurdles persist to serve outlying areas.

Steps are underway to fix connectivity problems, and a plan for e-learning is in the works. The plan highlights the role of tech tools to help everyone access education in Pakistan. The Senate group also stressed cheap internet devices' importance. They pushed the Federal Board of Revenue and the Pakistan Telecommunication Authority to find methods to involve all students in e-learning [9].

In this context, AI-powered bilingual conversational agent creates an exclusive chance to boost learning opportunities. This chatbot can tackle various hurdles in education, offering engaging, bilingual aid in Urdu and English. It's different from the usual e-learning resources, which typically cater to English-speakers only. A bilingual chat agent can involve learners who don't speak English, thus delivering educational info in a way that's inclusive, pertinent and wider-reaching for all. So, this project aims to create a bilingual chatbot for e-learning. It'll help students in Pakistan's less privileged areas, cracking the language barriers and education hurdles. It uses techs like large language model, natural language processing, and voice recognition to assist tailored learning. It fills the gap of language barriers and basic educational need by promoting e-learning for everyone.

1.2 Project Motivation

The motivation behind developing a bilingual conversational agent for e-learning in Pakistan was inspired by a real need. It's all about helping students from different language and economic backgrounds, to deal with tough basic education accessibility in far flung areas. In Pakistan over 26.2 million children out of school primarily in rural and underdeveloped areas, faces various issues to have access to basic education [8], one of which Language barrier such as English is one of the major hurdles, as many students, especially in rural areas, are more comfortable with Urdu than English. Also due to economic problems and minimal number of schools in country side the students are not able to get quality education.

The COVID-19 crisis spotlighted the need for different and alternative study resources. Students all over the nation depended on tools for online or distant-learning. But, most e-learning platforms in Pakistan mainly serve people who speak English. This might restrict Urdu-speaking students' access. An AI-powered bilingual conversational agent, fluent in both Urdu and English, has been tailored to fit Pakistan's language needs. This tool aims to break down language barriers in education, making digital learning tools easier for Urdu-speaking students to use. It offers bilingual support for a more engaging and accessible learning experience. This project centers around designing a user-friendly assistant that adjusts to personal language choices, boosting student involvement.

This chatbot fueled by AI speaks two languages. It uses technologies like natural language processing, large language models, and speech recognition. These help it chat meaningfully with students and lead them through their learning journey right as it happens. This addresses not only the language gap but also improves personalized learning as the agent can adapt responses based on individual queries and educational needs. So, by offering learning resources in Urdu, the chatbot matches the wider aim of including everyone in education. This is in line with the Pakistan Education Statistics (PES) report. For kids in areas with fewer chances, it's a big step up from the scarce on-ground tools. They get non-stop help for their studies.

This project is a leap towards better education in Pakistan. It's overcoming language and online education barriers. The chatbot interacts in two languages and guides on the spot. It interacts with students from diverse backgrounds. It provides them the resources they need to do well in their own language and English too.

1.3 Project Vision

We aim to build an AI-powered bilingual conversational agent for Pakistani students. The goal is equal access to online learning, no matter their language skills. Giving immediate help in both Urdu and English, this AI-powered bilingual conversational agent supports custom learning that includes all wide range of educational material. That way, it tackles a major issue in Pakistan's education system, access to basic education and language barrier. We want to do more than just provide educational materials. This conversational agent is designed to serve as an accessible learning assistant that students can use independently, enhancing both engagement and comprehension. By integrating speech recognition, natural language processing and large language model, the agent will allow students to communicate through voice in their preferred language, making it an inclusive solution for a diverse audience. By providing real-time feedback, explanations, and guidance on educational content, the agent will act as a personalized tutor, adapting its responses to meet

each student's unique needs and learning.

This project is seen as a lasting answer for improving Pakistan's online learning. It helps students with different languages background. Moreover, this aligns with Pakistan's goal for equal educational access. In the future, this tech might also support more languages or dialects further widening the range of audience.

The end goal is a learning space steered by AI, where every student no matter their command of English can shine academically. Through promoting a more all-embracing school experience, This project seeks to contribute to a more balance in learning community in Pakistan, where students from all backgrounds have the tools they need to excel in education.

1.4 Scope

Our goal? an AI-powered bilingual conversational agent. It uses the best AI tech, such as a Large Language Model. Its purpose? Help students learn more. It can work with spoken words or written text. This means it can understand and reply in either English or Urdu. With its wide information and data, it can help students of all ages, in all grades. Our solution is to provide an easy access to education. The platform will integrate the following core components:

Large Language Model (LLM): The Core and main component, an LLM will be employed to enhance the agent's capability to interpret complex queries and deliver accurate, contextually relevant responses across various subjects. The LLM will enable the system to adapt to diverse student questions by providing nuanced, subject-specific answers that support a broad educational curriculum.

Speech Recognition (STT): This unit works with live voice chats, letting students talk directly with the program. Thanks to speech recognition, the chat assistant can help those not as savvy with typing or texting, making it easier for everyone to use.

Natural Language Processing (NLP): By using NLP methods, the system can interpret and examine student questions, even if they're worded differently or are complicated. It will help the agent understand the true meaning of the questions and give useful answers. This boosts the student's understanding and interest.

Text-to-Speech (TTS): The TTS feature will change written replies into sounds you can hear in both Urdu and English. This helps students who learn better by listening or may have trouble reading. It makes the platform easier to use¬ no matter what language you speak or how well you can read.

User-Friendly Interface: The platform will have a simple, easy-to-use interface, good for students with different tech skills. This design puts users first. It's really

important so that all students, even those from rural or less developed areas, can easily find and use the features the agent offers.

Our main goal is to improve online learning for students in Pakistan's rural, less-served areas. Kids in these parts might struggle to get good educational stuff due to fewer schools and limited e-learning platforms. This handy chatbot in two languages aims to fix this. It gives a chat-based solution that can grow and adapt to add more languages and topics, if needed. Down the line, we may grow the platform to add more languages and top-notch learning features. Our big dream? To build a lasting, broad model for online learning in diverse areas lacking resources.

1.5 Problem Statement

This project tackles a significant issue. Pakistani rural students' scarce access to quality education, where language hurdles amplify their struggle with online learning. Shockingly, Pakistan Education Statistics 2021-22 reveal that 26.2 million children don't go to school [8]. Most of these children live in rural areas, where schools are scarce. Most online learning platforms today mainly use English. This makes it hard for students who don't speak English to be able to benefit from existing platforms.

Creating an AI chatbot that speaks both Urdu and English is the aim of this project. It's all about breaking down walls and offering custom learning experience. The service plans to boost inclusivity in education by giving prompt aid and speaking features, thus filling the gap of educational access in Pakistan.

1.6 Objectives

1. To create an AI-powered bilingual conversational agent using LLMs: Create a chat assistant using large language models (LLMs) that talks to students in Urdu and English or can chat via text prompts that provides educational help immediately.

2. To expand educational access: Give students in remote and less-developed places an e-learning platform they can use easily. That should provide educational help in Urdu as well as English. This will help overcome the issue of fewer schools in rural areas.

3. To utilize AI technologies for enhanced interactivity: Integrate high-level AI element such as, STT speech to text, natural language processing,TTS text to speech, text to speech and LLM large language model to create and Ai-powered conversational agent that makes learning more easy, engaging and

reachable.

4. To Support diverse academic topics across levels: Make sure the platform can handle questions from different school subjects. This way, it's flexible for all students, regardless of their studies or grades.

1.7 Tools

- Programming Language - Python: We picked Python as the number one language for our project. Why? It has a ton of features that we needed. Among them, it's great at machine learning (ML) and natural language processing (NLP) support.

- Embedding Model - Hugging Face Embedding: To translate the input text into dense representations of vectors, thus allowing the system to focus on semantic meaning and the context. Using pre-trained models with fine-tuning will produce greater accuracies for similarity matching, question answering, and sentiment analysis in the bilingual conversational agents.

- Text-to-Speech (TTS) Model for Speech Synthesis - Coqui TTS: Coqui TTS library, an open source advanced tool for producing human-like spoken responses in many languages, has been employed. The model offers flexibility so that phonetics can be adjusted both for English and Urdu and fine-tuned to be better.

- Speech-to-Text (STT) Model for Faster Transcription - Faster Whisper: This is an optimised version of OpenAI's Whisper model, optimised for high-performance speech-to-text transcription.

- Retrieval-Augmented Generation (RAG) - Langchain RAG: Integrates retrieval-based question answering capabilities to the system for the conversational agent to fetch relevant information from a predefined corpus or knowledge base.

- Large Language Model (LLM) - Meta Llama 3: Meta's Llama 3 is such a highly advanced large language model that can understand and generate human-like text in multiple languages. That makes the core of the conversational agent's capabilities.

- Flask - This is a Python web application with light back-end operations-managing API calls, processing input from users and server logic-which a lightweight library, Flask supports.

- JavaScript - To add interactivity, where it will be used for handling the client-side operations of the interface and handling the client-side functionality with the help of JavaScript, so that there will be real-time responses to the user as well as animations and the dynamic aspects of the interface are added.
- HTML - To include the standard structure of the web interface in order to ensure that the system is universally compatible with all browsers and devices. For input fields, buttons, and text displays, as well as other elements, to be defined better; hence make interaction possible between a user and agent in intuitive and accessible modes, HTML is essential.
- CSS - To use to style the layout and aesthetics of the user interface. The front-end will get a beautiful and consistent aesthetic appearance using CSS, and making the interface responsive to varying screen sizes.

1.8 Glossary

- **AI (Artificial Intelligence):** Is a computer science branch, making computers able to do things that usually need human brainpower which includes solving problems, making choices, and seeing patterns.
- **NLP (Natural Language Processing):** It is the ability of a computer program or software to understand human language as it's spoken and written, referred to as natural language. It's a component of artificial intelligence.
- **Speech-to-Text (STT):** A technology that translates voice into written text.
- **Text-to-Speech (TTS):** A technology that translates written text into speech.
- **Conversational Agent:** A virtual assistant or chatbot designed to interact with users through natural language processing and provide assistance or engaging in dialogue.
- **Bilingual:** Refers to the ability to in two languages.
- **LLM** (Large Language Model): A deep learning model trained on huge amount of datasets to understand and generate human-like response.
- **Corpus:** A large and structured set of texts or audio data used to train and evaluating large language models.

Chapter 2

Background Study

2.1 Background Study

Definition: A chatbot is sort of software application which enables to make it provide human-like conversation so that users can interact with digital systems in a conversational mode. Chatbots are utilizing natural language processing so as to interpret and respond to user inputs primarily through text or voice. These applications would be from simple rule-based systems, which react on the basis of keywords or phrases, to advanced bots that employ AI, driving through machine learning and deep learning to understand context, and personalize interactions based on the learning they get. Today, chatbots are being used in customer service, virtual assistance, healthcare, e-commerce, e-learning and many other domains where human-like conversation is valuable.

On a wide scale, chatbots can be broadly classified under several categories in terms of complexity and functionality:

1. Rule-Based Chatbots: These are simple chatbots that run on predetermined rules or scripts. They usually correspond to specific keywords or phrases and follow a decision tree to generate canned responses. They are not much flexible and cannot respond to complex queries beyond their programmed responses.

2. AI-Based Chatbots- These are essentially machine learning and natural language processing-driven chatbots. They can easily understand the intent of the user, context, and even learn over time. They are, therefore, more adaptive and capable of handling complex, open-ended queries, making them suitable for dynamic interactions.

3. Scripted or Guided Chatbots. These bots guide users through a channeled conversation, typically in customer service and support scenarios, by offering options to the user to click through to relevant solutions in a predetermined path.

4. Voice-Activated Chatbots: These chatbots, taking advantages from the voice recognition applications, can process and respond in the language itself. They are most typically found in virtual assistants and are often used in hands-free environments.

5. Transactional Chatbots: These bots are specifically tailored to the completion of specified transactions, like booking of appointments, processing orders, or making payments. The use of transactional chatbots is most popularly done in e-commerce and customer services for simple tasks.

6. Contextual Chatbots: With advanced AI, contextual chatbots can remember previous conversations and adapt to earlier conversations through some form of memory. This kind of memory will aid the bots in presenting a more personalized and contextual response to a user for a better experience.

This paper focuses on an AI-driven chatbot powered by a large language model (LLM). This would be based on a deep learning architecture, taking advantage of huge datasets, to generate highly coherent, context-dependent responses and therefore suited for handling complex conversations encompassing diverse topics with similar fluency as humans.

The development of these chatbots owes a lot to the LLMs, especially to OpenAI's GPT-series and other variants. These models assimilate very large amounts of text data and complex algorithms to learn linguistic patterns, grammar, and even context. Once they get a user input, these AI-led chatbots go on to analyze the text, decode intent, and come up with answers that are, quite often, nuanced and even contextually relevant. The new generation of chatbots can have flowing, multi-turn conversations, remembering earlier messages and fine-tuning responses to the need or requirement identified by the user. Not only do they respond in text, but most chatbots today use a combination of STT and TTS so that users can engage via voice commands and receive spoken responses, thus enabling greater accessibility and usability.

One of the key areas where chatbots are doing a lot of work is in customer service. Now, AI chatbots are found in many companies that aim to address routine customer questions, guide users through a process, and provide 24/7 support. This will definitely lead to increased user satisfaction and reduced waiting time. Take the example of the ability of a chatbot letting a customer track an order, reset a password, or troubleshoot a minor technical matter, thereby freeing up the time of a human agent for dealing with more complex issues. They are also helpful in e-commerce, where they make recommendations to consumers to help them decide in making purchasing decisions and answering questions about availability, pricing, or shipping.

Chatbots have also been adopted in educational institutions, especially AI-based ones. Chatbots play the role of a virtual coach for students in terms of answering their questions and walking them through the learning materials by providing responses that will fit the varying levels of understanding of the student. Bilingual and multilingual chatbots also broaden their utility in the sense that they can be utilized by individuals from different linguistic backgrounds to access learning content. This is very helpful for countries that have several languages spoken, in which case a bilingual chatbot fills the communication gap and offers support for learning in a more accessible manner.

Chatbots can also be developed to give information, guide users in pre-diagnosis, and help them access healthcare resources. In this context, AI-powered chatbots can prompt questions for the symptoms, offer explanations about the possibility, and guide the user on what to do next, such as get a doctor, among others. While these chatbots are certainly no replacement for professional medical advice, they do possess value as a preliminary support system that could be used to better understand medical concerns.

The development of chatbots powered by advanced AI models is a giant step forward in the evolution of digital interaction. Today's chatbot is far beyond a simple question-answering tool, but virtual assistance which nowadays ensures more efficiency, personalized support, and enhancement of user experience within different industries. The addition of voice capabilities, bilingual options, and deep-learning algorithms pushed the development of this technology even further, so more people can attain greater accessibility and context awareness in interactions. With the development of AI technology, chatbots shall become more intuitive and important constituents of many digital ecosystems, ranging from customer service to education and even more.

This paper discusses how to design an AI-based chatbot that may prove helpful for e-learners while seeking apt materials relevant to their educational needs. Firstly, the designing principle of this paper aims at creating a chatbot that could communicate with students in Urdu and English and extend educational material recommendations based on learners' requirements.

Mostly, in education settings, it has been seen that the trend of applying a chatbot is rising to address queries from students and to help learners resolve their questions while learning. Many studies have used rule-based systems, coupled with the integration of A.I. technologies, to develop these kinds of chatbots for use in meeting the demands of the users. However, such studies mainly focus on a monolingual page or a question-and-answer system that must be put more on neural engineering and N.L.P. This paper discusses the development of a retrieval-based A.I. chat box for a student trying to get answers to teacher-based questions from a

specific database. It is used in multilingual contexts.

Humans have always wanted to create anything that can listen to and respond to them as human creations. In this regard, the first person who conceptualized the chatbot was Alan Turing, who did so in 1950. He asked an interesting question, "Do machines think?" Turing's scenario, which describes the behavior of an intelligent machine, is precisely what most people refer to as a chatbot.

Artificial Intelligence becomes more intertwined into our daily lives through the development and examination of intelligent software and hardware, which is termed as intelligent agents. Intelligent agents can be programmed to perform a wide range of jobs, from labor work to sophisticated operations. A chatbot is a typical example of an AI system and one of the most elementary and widespread examples of intelligent Human-Computer Interaction (HCI). It is a computer program that acts like an intelligent entity when conversed with by text or voice and understands one or more human languages through Natural Language Processing. According to the lexicon, a chatbot is "A computer program designed to simulate conversation with human users, especially over the Internet". More popular terms for chatbot include a smart bot, an interactive agent, a digital assistant, or an artificial conversation entity [10].

Why?

Why do the users use the chatbots? It seems that the chatbots really provide great potential for swift and handy support specifically answering questions of the users. The productivity is considered to be most frequent motivation for users who employ chatbots. Other motives are entertainment, social factors, and contact with novelty. However, to balance the above-mentioned motivations, a chatbot needs to be designed so that it can work both as a tool, as well as a toy and a friend [10].

After the COVID-19 pandemic and its subsequent lockdown, teachers and learners are well-familiarized with alternative ways like distant teaching, assessment, and grading. Lately, there has been a growing demand for workable alternative study methods resulting in adopting sophisticated educational resources like chatbots [11].

Market Research Future, the company that analyses markets, definitely supports this view: In the "coming seven years, we should anticipate dramatic progress, especially in conversational artificial intelligence market including conversation chatbots". Forecast Suggests the volume of the market might reach USD 32.5 billion, growing at a compound annual growth rate of 22.6 percent from 2022 to 2030 [12]. This prediction is consistent with other studies issued by a market research specializing firm - Markets and Markets- that foretells massive growth for this industry during the coming five years. The market size is expected to grow from USD10.7 billion in 2023 to USD29.8 billion by 2028 at a time when the average CAGR will

hit almost 22.6 percent [13].

The new intelligent chatbots have emerged due to past achievements by technology concerning the large language Models (LLM). These kinds of chatbots can now understand users' intentions, converse in a manner of ordinary complex where the user interacts with them quite rapidly and efficiently, and deliver relevant cost-effective assistance and support to users [14, 15]. With the advancements in Large Language Model (LLM) Generations of smart chatbots have been developed and it has become possible to understand the user intent, conduct conversations close to natural ones, and provide the user with customized and affordable interactive assistance and consultation [14, 15].

According to Abashev (2017), "Chatbots are artificial intelligence tools that can already communicate with users, though their development has become much farther since the creation of the first one, E.L.I.Z.A. It is more sophisticated with the modern chatbots of today, like Siri, Alexa, and Google Assistant who tap into A.I. to hold on to perfectly random spontaneous conversations while completing various chores.". With calorific burnt achievements in machine learning, it has grown exceptionally robust and said even to be able to grasp emotions [16].

Chatbots are also more and more integrated into schools as a means to raise important questions of students and support the process of learning. Researchers have created many rule-based or A.I. systems, able to give fast and proper information to the user. However, most studies only one language or only simple ask-and-answer commands. Fewer in number investigate using neural networks and Natural Language Processing (N.L.P.) methods in multilingual settings [17].

Education sector chatbots interact with students and answer their queries. They provide evaluations and feedback for the learners and even offer communication and coordination services to students and teachers. There are many benefits to teachers that can reduce some tedious tasks, for example, smart chatbots who can follow attendance, create presentations and tests, distribute papers to mark grades along with marking the assignments and the tests. Among the students' perimeter of learning within the university, these chatbots answer queries by the students give ratings, write feedback, and help student-teacher interaction and collaboration in the education process. There are many benefits to educators, and smart chatbots may reduce the time involved in the most time-consuming processes of taking attendance of students, preparing and delivering examinations and class work, as well as preparation of class work and examination scores [18].

The Bandore was developed by Professor Ron Lee for the purpose of helping students enhance their usage in English. It was the winner of 2018 Loebner Prize Contest (Ferrara, 2016). A kind of Turing Test, the Bandore is used to test the existence of A.I [19]. As Park (2019) said, this system enables the simulation of

changes in spoken languages and deep knowledge about interconnected themes of languages, culture, politics, history, and geography. Wang, however, presented it to be somewhat sweeter in tone than most systems, and that could lead to more significant usage of the system [19]. In 2007, Abu Shawar and Atwell realized that an area of application for chatbots also includes foreign language learning. They propose a different model for the interaction which has been applied so dominantly, that of a student-pupil relationship. An example of such a chatbot is the Tutor Mike chatbot.

History

Artificial Intelligence has a rich history and one that spans several decades of great innovation and progress. The idea of AI actually goes as far back as to some ancient myths and stories relating to mechanical beings capable of intelligence. However, from the formal foundations, AI has a 20th-century underpinning, thanks to British mathematician and logician Alan Turing [20, 21].

In 1936 Turing introduced his paper on the concept of computation that introduced the "Turing machine", a theoretical device that became fundamental to understanding what could be computed [22]. In 1950 he published Computing Machinery and Intelligence, containing the famous "Turing Test" - a method for determining whether a machine could exhibit intelligent behaviour indistinguishable from that of a human [23]. Turing's work serves as the foundation for the first wave of AI research and provides a ground for intelligent conversational agents, or chatbots, which will later come to be an accessible way of testing machine intelligence in dialogue [21].

Major concerns in early AI research during the 1950s and 1960s were symbolic AI, involving programming machines for manipulation of symbols and logical reasoning. The Dartmouth Conference of 1956 led by John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon is usually considered AI as an area of academic study. This conference settles the nomenclature and coins the term "artificial intelligence" and sets very ambitious goals for machines to be able to have actual human-like intelligence [24]. Early work in AI has involved some of the world's most renowned foundational work, such as building systems like Logic Theorist developed in 1955, which proved many mathematical theorems and ELIZA, one of the world's first chatbots, developed in 1966 [25].

Artificial intelligence experienced periods of boom and bust. The term "AI winter" was coined for the period of the 1970s when funding and interest in AI research declined as people became disappointed over unsatisfactory results. Lack of practical effects, coupled with the capabilities of computers of those times, had led to a drop in enthusiasm. This notwithstanding, the more specialized version of AI-and for instance, expert systems that imitated the decision-making capabilities

of human experts in specific areas-kept evolving even in the 1980s [26]. Chatbots, though, continued to be simplistic rule-based systems with less functionality because of the technology at the time [27].

The rebirth of AI in the late 1990s and early 2000s had their basis in developing machine learning, perhaps more so with algorithms that enabled a computer to learn from data. A new milestone was reached in 1997 when IBM's Deep Blue defeated world chess champion Garry Kasparov in a highly anticipated match, demonstrating that AI could be used to surpass an expert's best performance in various complex problem-solving tasks [28]. This was the era that saw enhanced ways of handling NLP, where it gradually paved its way for the more complex chatbot performance and eventually goes beyond simple preset word answers [21].

Big data explosion, increased computational capabilities, and deep-learning breakthroughs helped build rapid advances in AI in the 2010s, primarily in the technologies of natural language processing and computer vision [29, 30]. AI was started to begin outperforming humans in few specific tasks such as image recognition, translation, and playing strategic games like Go. Chatbots also grew strong with the technologies of neural networks and deep learning, allowing them to learn from large amounts of data and spawn more accurate and context-aware responses. Such was the era that heralded the emergence of virtual assistants such as Siri, Alexa, and Google Assistant, bringing chatbots into people's everyday lives in the shape of interactive tools to communicate the power of AI to users [31, 32].

Previous Work

The production of E.L.I.Z.A. in 1966 is generally regarded as the first prototype of chatbot technology. Joseph Weizenbaum developed E.L.I.Z.A., an early concept for natural language processing based on a keyword-match method in identifying appropriate parts of a user's input in response. Architecture of the chatbot consisted in application of control and patterns of language to simulate a conversation by reflecting inputs to the user in phrased or guided form. This allowed E.L.I.Z.A. to mimic the interaction style of a psychotherapist, which marked one of the earliest examples of human-computer dialogue. Even though this was a major breakthrough, the limitations of E.L.I.Z.A. were quite evident: its database and conversational flexibility were restricted due to poor contextual understanding as well as fixed pattern matching. Thus, E.L.I.Z.A. paved a path to subsequent experiments but proved not to be very flexible to new contexts-few enhancements were achieved in the 1980 with the birth of A.L.I.C.E. (Artificial Linguistic Internet Computer Entity), a chatterbot that introduced artificial intelligence markup language (A.I.M.L.) for better understanding the conversation on a dialogue level. A.I.M.L. is an extension using XML and has features designed that would enable the chatbot to realize and respond in more dynamic manners to the patterns of a

conversational process. A.I.M.L. followed the organization of knowledge in a tree-like topic, class, and category hierarchy, with categories being the base structure for structuring rule-based dialogues. Each pattern of rules determines the type of user input it will recognize, and the templates of rules gave A.L.I.C.E. proper output capabilities. This kind of structure gave a modular and more expanding knowledge base. A new A.I.M.L. object was easily added, and the pattern-matching abilities of the chatbot were honed even further. A.L.I.C.E. showed an important development in the field of the chatbot, especially in respect to structuring and expanding of knowledge bases [33].

In continuation, ChatScript was a direct result also, an evolution from A.I.M.L., some more mature mechanics in the concept of understanding, as well as response generation. Known for powering chatbots that later won the Loebner Prize, this was intended to handle user input in the form of semantic analysis, digging much deeper and getting more contextually correct responses. This technology marked a change from simple pattern matching towards a method that connected user inputs with a specific topic, an entity, or context applying dedicated rules to each area of knowledge. Thus, correlation of inputs into specialized categories enabled the chatbots to generate much more accurate and semantically rich answers. Meaning-focused innovation over pattern recognition constituted a critical advancement in conversational AI [33].

PARRY was created in 1972 by American psychiatrist Kenneth Colby as a program that could emulate the behavior and thinking of a patient with schizophrenia to try to replicate the cognitive processes associated with the disorder. It is an NLP program and is designed to represent the confused, sometimes illogical thinking reported in people afflicted with schizophrenia. In other words, PARRY operated using a sophisticated system of assumptions, attributions, and emotional responses as it underwent transformations due to varying weights assigned to verbal inputs. The responses were predetermined to mimic and mimic the responses and thinking styles typically found in schizophrenic patients. To determine the efficacy of PARRY, it was tested according to a version of the Turing Test. Very early in 1970, human interrogators, communicating with the program by remote keyboard, could not distinguish PARRY from a real person exhibiting unreasonable or schizophrenic behavior; they couldn't tell, as a matter of fact, any more often than not, with accuracy no better than random [34].

Another of the most well-known 1980s chatbots is Racter, a software program written in 1983 by William Chamberlain and Thomas Etter and published by Inrac Corporation. The term is coined from the word "raconteur," which refers to a storyteller: it was an intentional AI text generator. Chamberlain explains in the introduction to The Policeman's Beard is Half Constructed the special function

of Racter's text generation: "An important faculty of the program is its ability to instruct the machine to choose variables (words or phrases) randomly, and the machine would then retrieve them at some point and include them in whichever block of text it was generating" [35]. This feature made Racter produce texts that, though at first glance seemingly random, contained a form of "order," thus imparting coherence along with meaning to the output. "Whereas the text might be nonsense, it was still a concept and even evoked an odd kind of beauty in the text, showing that, as Chamberlain suggests, "madness" can be "thinking" if there is beautiful English involved" [36]. Racter is written in BASIC for a Z80 microcomputer with only 64K RAM-an early system by any measure of computing power. Its capacity is usually less than what would be needed to create complex texts even on some of the most primitive systems. The prose of Racter has been labelled an a "poetic rhetoric", short text snippets that seem to almost reach an avant-garde style where the AI-generated prose goes for glimpses in human-like expression. The question comes up here: is prose thought of as being actually prose if it originated without human intelligence? Should our judgment of prose extend to include art created by AI, that-lacking conscious human intent-may yet express meaning in novel forms?

Jabberwacky in 1988 (Jabberwacky, 2019). In turn, Jabberwacky was written in CleverScript, a language relying on a spreadsheet basis for creating chatbots easily, and it relies on contextual pattern matching to react based on previous talks. However, Jabberwacky cannot answer at a high speed and work with a huge number of users (Jwala, 2019) [37].

In 1991, another new chatbot was made based on innovative and revolutionary technology. The innovation was the Sound Blaster sound card; which Five Labs had come up with. This was yet another landmark in the history of the development of chatbots. Dr. Sbaitso was the name of the application, officially called "Sound Blaster Artificial Intelligent Text to Speech Operator." It is [38]. "it was capable of synthesizing speech, which is making sounds that convey language in a certain way that let it be more human than its earlier versions, though it could not speak in a more complicated and advanced way either. This problem, that is that chatbots have less depth in their speech than they should be, is still the case.

In 2001, success seemed to be knocking at the door of a new step in its evolution as SmarterChild (Molnár and Zoltán, 2018). The bot was accessible on Messengers, America Online, and Microsoft M.S.N., among others. It was the first case the capability of the chatbot to make people cope with their everyday actual tasks. It could ask such information as movie programs, the outcomes of sports games, stock exchanges, news, and weather. It was only by the intercommunication with the chatbot that I could get into information systems, and it opened a significant evolution in trajectories both in machine intelligence and human-computer interaction

but as indicated by Indra marked that this had limitations.

I.B.M. developed the new A.I. chatbot named Watson in 2011 (Watson Assistant I.B.M. Cloud, 2020). Known answers and unknown questions were the quiz game "Jeopardy" format played by two previous title holders. Not only was he an expert in jeopardy, but he could also understand humans better and won both titles. Years have passed, but Watson has made it possible to empower businesses with more powerful virtual assistants. Watson Health was designed for the diagnosis of disorders by health practitioners. One of the major drawbacks for Watson is that it can only converse in the English language [14].

Recent Developments

Deep learning and the neural networks brought the emergence of chatbot technology. Natural language can now efficiently be treated for accuracy and flexibility in response with deep learning. Such convincing agents that have been developed are like Xiaoice by Microsoft in China, which aims to create a sense of relationship through the generation of emotions with appropriate responses. It became one of the most popular bots in the entire world with millions of users and was used as a model for future emotional and social AI-based products.

That same year, Facebook Messenger launched business chatbots that allowed companies to provide automated customer support, show users personalized recommendations, and process simple transactions. Based on this innovation by Facebook, more companies began to use chatbots in e-commerce to minimize the hassle that customers go through trying to get help from the company and involve themselves with the product.

OpenAI set the new benchmark in the possibilities of chatbots for 2020 with its mammoth language model of 175 billion parameters known as GPT-3 (Generative Pre-trained Transformer 3). GPT-3 can talk fluently on a wide variety of topics without an explicit instruction to do so, hence the display of unparalleled generalization abilities. The transformer-based model so strongly enhanced the conversational ability of chatbots that it was optimized to adapt to different topics as well as understand nuanced input from language. The successor, GPT-4 was in 2023 developed and took the previous capabilities to a new level by enhancing contextual understanding, a multilingual capability, and even handling complex conversational tasks better.

More recently, voice-activated assistants like Amazon Alexa, Apple Siri, and Google Assistant have tapped on very advanced NLP models. These assistants have undertaken so many different kinds of tasks-from reminding users to controlling smart home appliances to complex queries-making their way into the routine of everyday life as conversational AI.

2.2 Limitations

Chatbots have dramatically changed the type of interaction, customer service, and many other elements related to digitization. Though much so far has been achieved through the means of chatbots, there are yet many constraints the effectiveness and reliability of which they can hardly cross. These include the following:

Limited Context Understanding: The most critical flaw of most chatbots, especially rule-based ones, is their poor capability to perceive the context or nuances involved in a discussion. AI-based chatbots that are based on NLP have further capabilities but possibly still cannot interpret ambiguity, sarcasm, or informal text and might instead lead to confusion or produce an inappropriate response. That said, it could cause frustration for the user because the chatbot cannot pursue a complex and contextualized conversation like a human would [39], [40].

Issues Solving Complex Questions: Simple queries can be dealt with; however, chatbots choke on intricate or multi-step problems. Based bots and even some AI-based bots without advanced training are not able to manage complex questions or multi-faceted inquiries as much as others and hence may turn out to be less useful when applications require deep problem-solving or reasoning. Customer service applications often face numerous issues that can be relatively simple yet can also be complications in processing [41], [42].

Dependence on Quality and Quantity of Training Data: AI-based chatbots largely depend on the relevance and diversity of the training data through which their model is built. Therefore, in case these data are biased, outdated, or limited, the chatbot will produce biased or irrelevant responses. It will further lead to a lack of cultural sensitivity or failure to accustom to different language patterns so that diverse audiences across different regions cannot use the functionality of the chatbot effectively [43], [44].

They cannot be said to demonstrate real emotional intelligence. A hugely differential lies between the possible 'intellect' of a chatbot which is well-developed and driven by artificial intelligence and human 'emotional understanding or intellect,' someone who could empathize with your feelings. They could be programmed to parrot empathy words with set phrases, but without the deciphering capacity and responding to your emotional cues, responses can sound like done by a preprogrammed automaton or robot rather than an intelligent-thinking human being. This lack of emotional intelligence is particularly limiting when the job entails customer support or therapeutic communication, as these roles depend on the interactions flowing from the chatbot being inherently empty or impersonal [45], [46].

Security and Privacy: Generally, chatbots collect and process a lot of user data; hence, they can pose significant security and privacy concerns. Otherwise,

without proper measures in place for data protection, the resulting chatbots can easily become weak spots for cyber attacks or inadvertently reveal sensitive information. This will be particularly important if the industries concerned happen to be regulated, such as finance or healthcare, where inappropriate data handling leads to legal and reputational issues [40], [47].

These limitations depict areas that chatbots with simpler architectures need to be improved significantly-and indeed all of them. Even though AI-based chatbots that have large language models possess strong capabilities, research continues toward removing these constraints and will permit more human-like, secure, and contextually accurate interactions with chatbots [46], [42].

A look at table 2.1 mentions the flaws of several of these chatbots since they belonged to different generations. Each was revolutionary at its time but had some limitations that restricted its operation and user experience. ELIZA faced difficulties pertaining to context as it could function only on keyword matches while A.L.I.C.E., though it introduced AIML for conversational dialogues, was still limited by its rule-based, static structure. Later versions, like PARRY and Racter, emphasize further development toward an ability to mimic a precise mental state or generate creative prose but are narrowly confined by a lack of larger conversational adaptability and learning ability. Even the modern chatbots, including Google Assistant and Amazon Alexa, which are substantially more sophisticated than older systems, rely heavily on their parent ecosystems and sometimes get contextual interpretation or voice recognition wrong. Such a development sequence reveals that though chatbot technology has indeed come a long way, each generation has been plagued by inherent problems, hence the need for constant innovation in this discipline.

Previously Developed Chatbots: For a comparison of the limitations of previous chatbot systems, refer to (Table 2.1).

Table 2.1: Comparison Table of Chatbots

S#	Chatbot Name	Limitation
1	ELIZA (1966)	Lack of Emotional Intelligence
2	PARRY (1972)	Lack of Learning Ability
3	RACTER (1983)	No Understanding of Context
4	ALICE (1995)	Outdated Technology Framework
5	JABBERWACKY (1997)	Limited Scalability
6	SIRI (2011)	Dependence on Apple Ecosystem
7	WATSON (2011)	Complexity and Cost
8	ALEXA (2014)	Inconsistent Voice Recognition
9	GOOGLE ASSISTANT (2016)	Dependence on Google Ecosystem
10	TUTOR MIKE (2017)	Narrow Knowledge Base

Chapter 3

Methodology

This chapter explains the methodology followed to create the AI-powered bilingual conversational agent. It further describes the design, architecture of the system, and all the processes that had to be undertaken to implement it.

3.1 System Requirements

Functional requirements include the functionalities that must be supported by a chatbot to create effective, bilingual interaction. Here are the detailed aspects:

1. Language Choice: It can also allow a user to be able to choose his preferred language at the start of the conversation.
2. Language Detection: The chatbot should, therefore, automatically identify the language of the input if no language is stated by the user in the input.
3. Response Generation: The chatbot should generate responses in the same language as the user's input.
4. Context Maintenance: The chatbot must track the context of the conversation within a session, allowing it to understand follow-up questions or statements in either language.
5. Bilingual Knowledge Base: The knowledge base of the chatbot should support content in both languages.
6. Exception Handling: If the chatbot does not recognize the input or is unable to produce a relevant response, it may suggest reformulating.

Non-Functional requirements talk about how the system is able to perform its functions and overall experience by the user. Here's a detailed view:

1. Performance: The average response time needs to be less than 15 seconds for 90 percent of interactions.

2. Usability: User Interface shall be designed so that it is easy to use.

3. Accessibility: The chatbot should be accessible to users with disabilities.

4. Scalability: The architecture of the chatbot should support future expansion, such as adding more languages or features without significant redesign.

5. Availability and Reliability: Ensure the availability of the chatbot 90 percent of the time because the users need it 24/7.

3.2 Architecture And Design

System architecture and design refer to the overall structure, components, interfaces, and attributes of a system to satisfy requirements and constraints set by stakeholders. This process aims at breaking down a system into its most important components, analyzing the relationships and interactions among these components, and determining how they will work to produce the intended result of the system.

System architecture and design aim to create an adaptable, scalable, and low-maintenance system that meets the needs of users and other stakeholders, and can be built and operated with less effort and cost.The architectural and design schematics captured in this section illustrate these objectives and principles in action.

1. **Flow Chart**: A graphical representation of a process or system, where standardized symbols and arrows define the sequence of steps involved in the process or system. A flowchart visually maps out a task or workflow, making it easier to understand, analyze, and communicate the complexity of processes involved. Figure 3.1 shows the flowchart of an AI-powered bilingual conversational agent.

2. **Architecture**: Architecture refers to a structured design and organization of a system, its components, the interrelationship among them, and the manner in which they interact to achieve certain objectives. In giving blueprints for designing, managing, and evolving systems, architecture must meet functional and nonfunctional requirements. Figure 3.2 shows the overall architecture of the chatbot.

3. **Use Case**: A specification that comprehensively describes how a system or application will interact with users (or another system) to achieve a certain goal. It outlines the steps of interaction between the user (actor) and the system, from initiation to completion, for a clear understanding of the functionality the system provides. Figure 3.3 shows the use-case diagram of the chatbot.

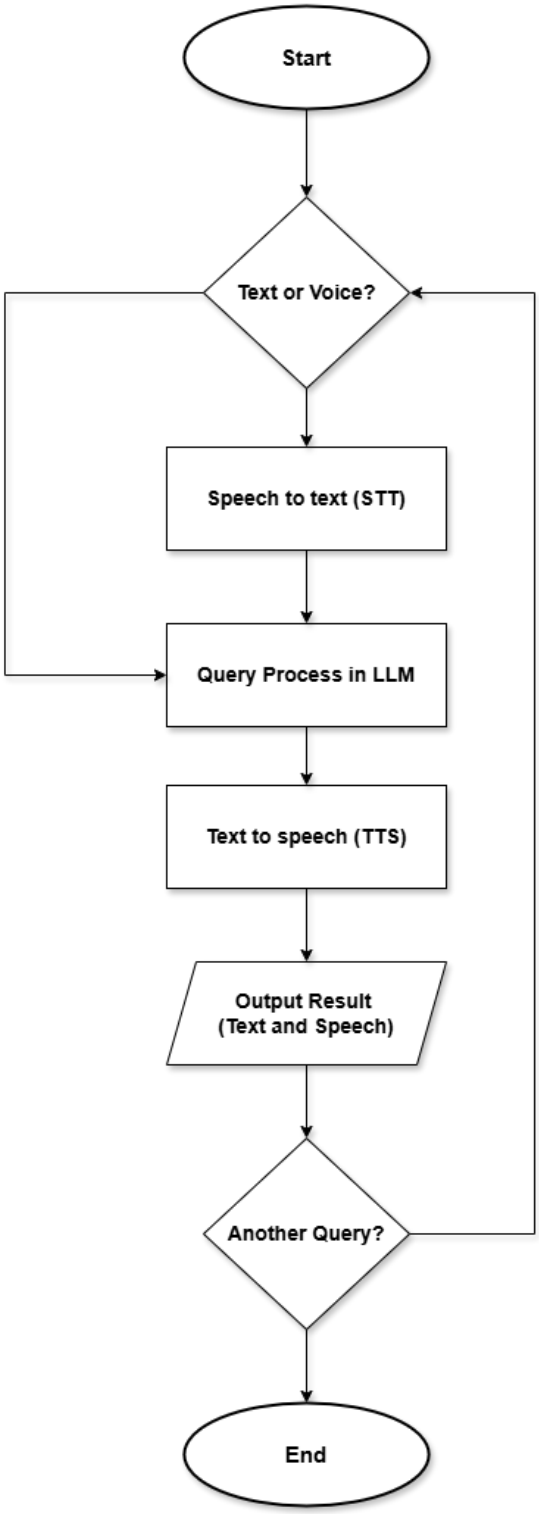

Figure 3.1: Chatbot flowchart

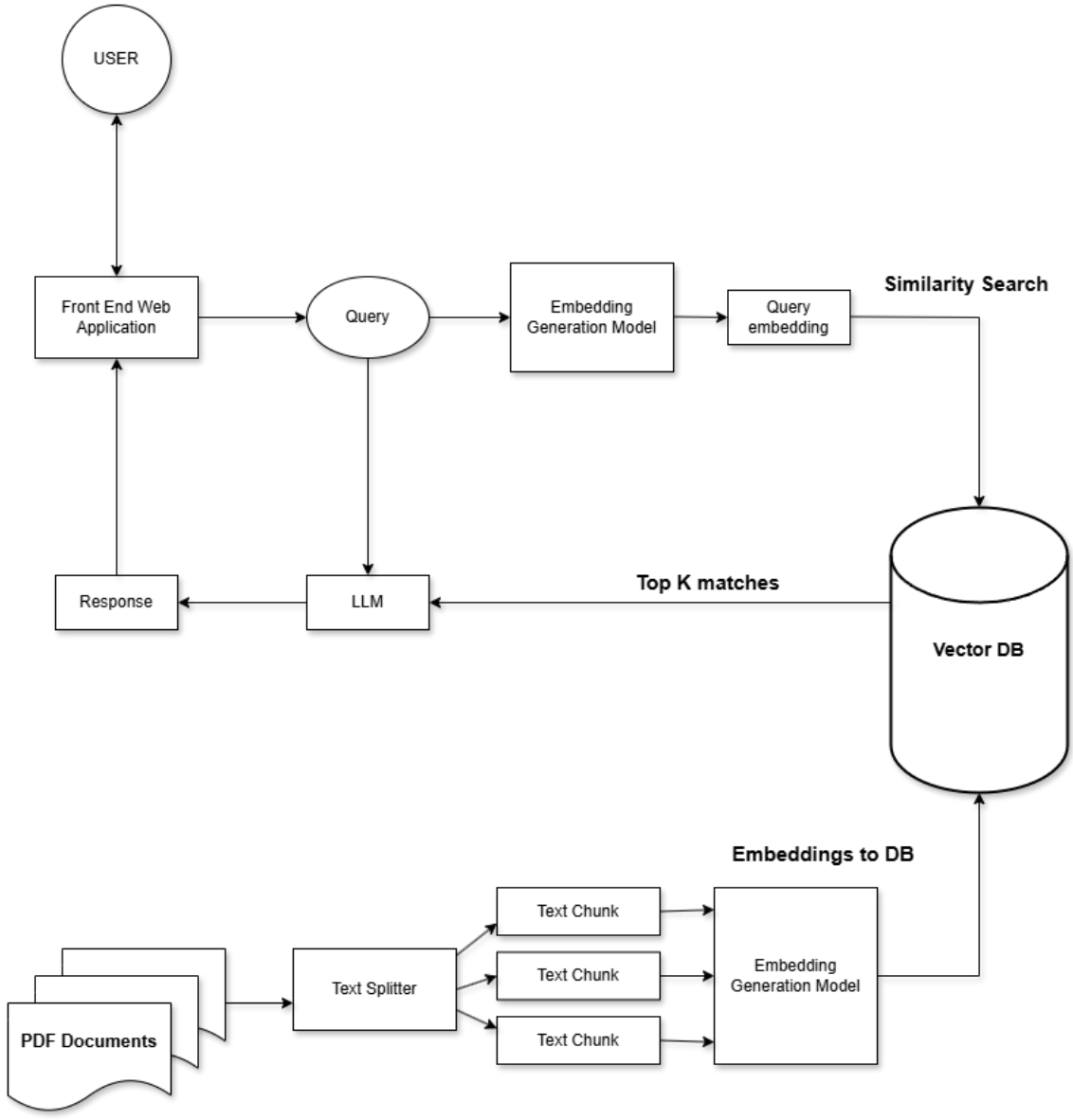

Figure 3.2: System Architecture

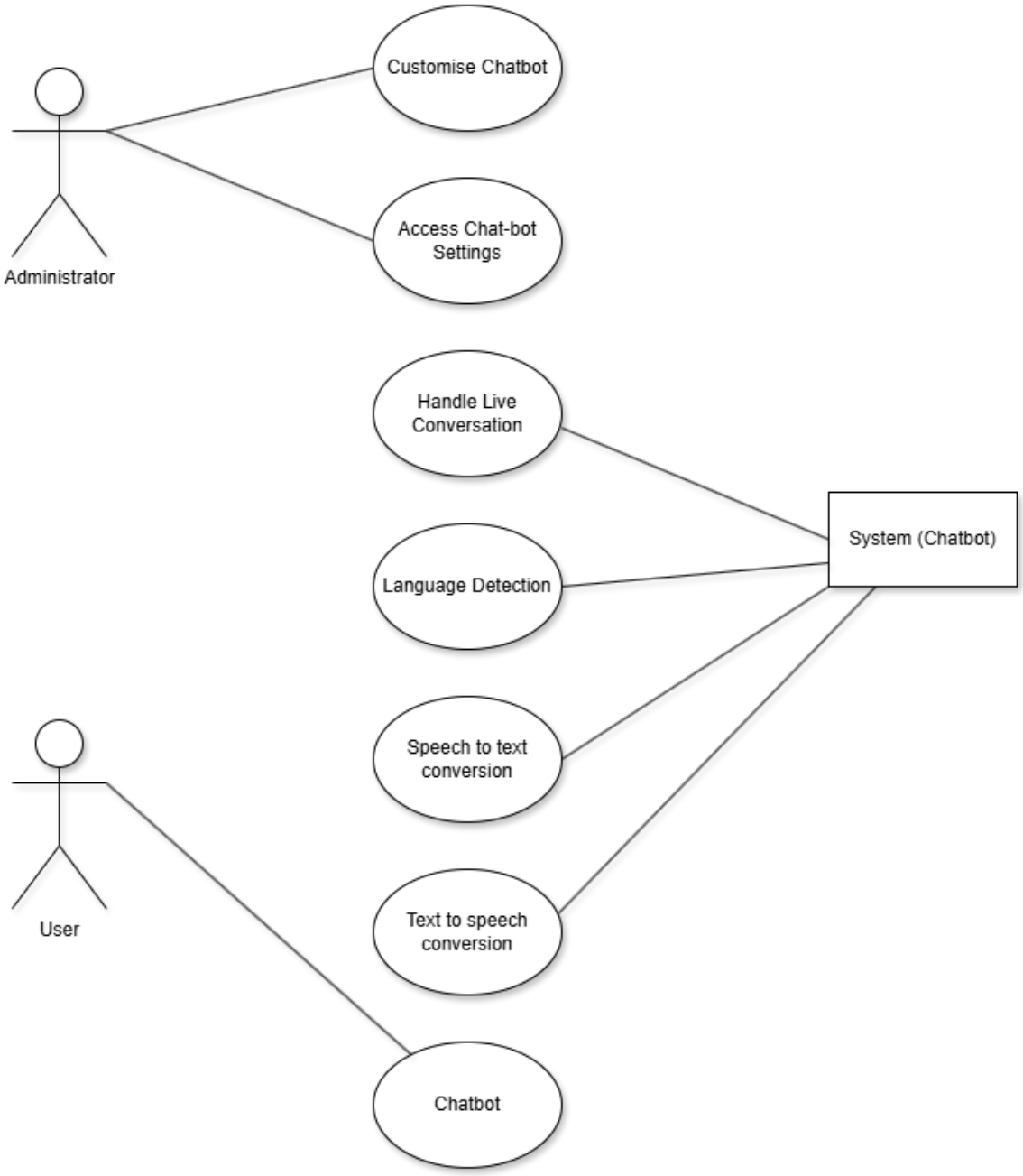

Figure 3.3: Use Case Diagram

4. **Expanded Use Cases:** The tables below describe various use cases of the chatbot in different scenarios.

Use Case 1

Name	Chatbot Query Handling
Actor	Chatbot User
Purpose	Provide accurate responses to user queries
Description	This use case involves the chatbot processing the user's query and generating a response, regardless of the input method. It describes how the chatbot interprets and responds to queries after receiving the input.
Pre-Conditions	1) The chatbot is operational and running, 2) The user has internet connectivity, 3) The chatbot has access to relevant data or APIs for processing the query.

Use Case 2

Name	Input Mode Determination
Actor	Chatbot User
Purpose	Detect whether the user input is text or voice and process it accordingly.
Description	This use case focuses on detecting whether the user's input is in text or voice format and ensuring the system processes it correctly (e.g., converting voice to text when necessary).
Pre-Conditions	1) The chatbot is operational and running, 2) The user has internet connectivity, 3) The device being used supports both text and voice input modes.

Use Case 3

Name	Provide Fallback Responses
Actor	Chatbot User
Purpose	Handle situations where the chatbot cannot understand the input or provide a specific answer.
Description	When the chatbot is unable to understand the user's input or lacks the information to generate a meaningful response, it provides a fallback reply and suggests alternative actions.
Pre-Conditions	1) Chatbot must be operational but may lack sufficient data for a specific query, 2) User must have internet connectivity.

5. **Data Flow Diagram**: It brings out a graphical representation that describes the flow of data through a system, whereby inputs are transformed into outputs through various processes. From this, one may draw a high-level view about the components of the system, the sources of the data, and how things interact. Figure 3.4 shows the Data Flow of the chatbot.

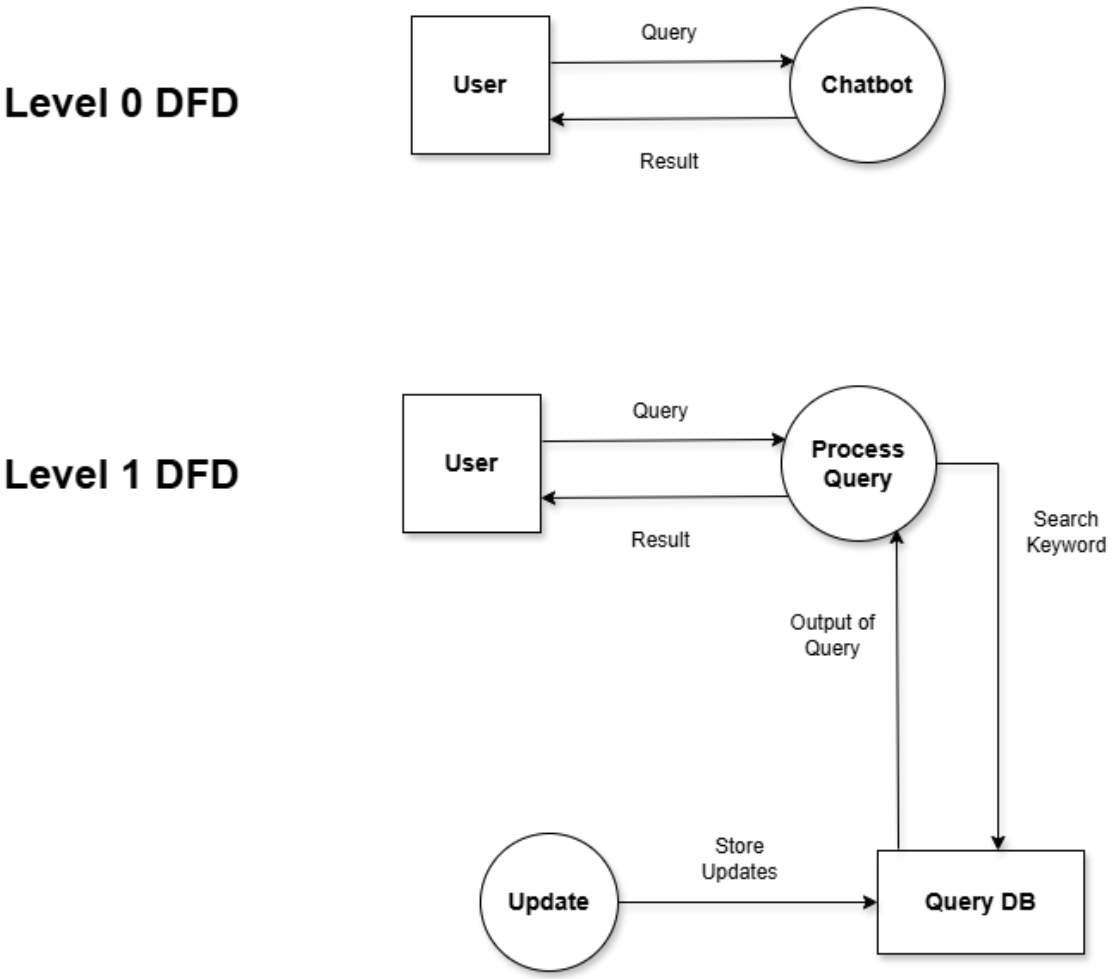

Figure 3.4: Data Flow Diagram

3.3 Features Description

Features of the Chatbot System The chatbot supports bilingual, culturally inclusive experiences using intelligent voice-based interaction, proficient in both Urdu and English languages. Here are its key features:

Real-Time Voice Interaction

- Voice Capture: It captures live audio from the user by using a microphone.
- Voice Detection: It detects when the user starts and stops talking for smooth conversations.
- Bilingual Support: Can listen and process both Urdu and English fluently.

Speech-to-text transcription

- High Accuracy Transcription: High accuracy transcription uses the faster whisper library to transcribe audio files into text in both Urdu and English.
- GPU-Accelerated: It makes transcription fast as it allows real-time interaction.
- Urdu and English: specific models optimized for speech recognition.

Context-Aware Answers

- AI-Driven: Answers Powered by MetaAI to generate dynamic, contextually relevant responses in either Urdu or English based on the input of users.
- Cultural Relevancy: Written in the parlance of Urdu-speaking as well as English-speaking languages.
- RAG, or Retrieval-Augmented Generation: uses Pinecone in retrieving relevant data to generate factual and enriched responses.

Text-to-Speech Conversion

- Bilingual Speech Synthesis: Using Coqui TTS for the synthesis of natural speech in Urdu and English.
- Clear Pronunciation: Produces pronunciation suitable for bilingual dialogue with clear and distinctive accents.

Users Web Interface

- Front-end Tools: Developed with HTML, CSS, and JavaScript using a simple and interactive design.

- Audio Controls: Listens to synthesized Urdu or English response with some playback options
- Real-time Feedback: Transcribed text and synthesized audio are provided immediately.

Bilingual Support

- Dual Language Support: Especially for customers who prefer to write in Urdu, English.
- Scripts Recognition: It will accept input in Roman Urdu- written text in Urdu using an English alphabet-on a keyboard or pure input in English.

Noise Handling

- Filter out the background noises: Deletes background noises so that the voice can be correctly detected.

Scalable and Modular Framework

- Could run on Local or Cloud-based setup: It supports personal use as well as scaling up to a great extent on AWS.

3.4 Data Preprocessing

Audio Preprocessing of Data

- Noise Reduction: Audio Filtering Techniques to remove background noise.
- Normalization: Adjust audio levels to maintain consistency in volume across the dataset.
- Segmentation: Divide a long audio recording into small pieces using silence detection.
- Sampling rate: Resample audio data to 16 kHz, which is optimal for transcription models like Whisper.
- File Format: Save file as .wav. This file type is accepted by almost all speech processing libraries.

3.5 Libraries

The project's implementation will entail integrating both the backend using various Python libraries and modules, as well as frontend technologies that help build AI-powered Chatbot for e-learning. The libraries and tools are meant to provide functionalities like audio processing, transcription, text generation, synthesis of text-to-speech, semantic understanding, and, finally, design an interface that is user-friendly.

Libraries:

1. Os: Controls environment variables.
2. Collections: It controls audio buffers through deque, which is so efficient in real-time operations.
3. Numpy: Takes raw audio data and also computed sound levels to initiate voice detection.
4. Wave: Saves the recorded audio to .wav format for further transcription.
5. Pyaudio library: Enables recording real-time audio from the computer's microphone.
6. Faster-whisper: This is an optimized version of the Open-AI Whisper model, which accelerates transcription processes, speech to text.
7. Torch.cuda: Checks for available Gpus to accelerate the Whisper model.
8. Coqui-tts: Convert any given text to high quality speech by deep learning models.
9. Meta-ai-api: Interact with the Meta AI model llama 3 to develop AI-driven responses that can really be conversational.
10. Langchain: Chain several calls of the model and tools for better AI responses.
11. Hugging Face Embedding Models. Provide semantic understanding and context matching for better conversation quality.
12. Pinecone: A vector database is used for efficient similarity search in RAG, considering getting context-based retrieval of information that is relevant.
13. Flask: The back-end logic or API endpoints to serve requests coming through the web interface.

3.6 Hardware Description

Hardware Requirements for the AI-powered Bilingual conversational agent for e-learning.

Minimum System Requirements

- Processor: Quad-core CPU (Intel i5 4th gen or AMD Ryzen 3 1300x).
- RAM: 8 GB (sufficient for small-scale usage).
- Storage: 20 GB free disk space (for dependencies and temporary files).
- GPU: Not required; CPU-only operation is possible for lightweight tasks.
- Audio Setup: Standard microphone and speakers or headphones for voice input/output.
- Internet Connection: Minimum 5 Mbps consistent connection to prevent latency in responses.

Recommended Hardware Requirements

- Processor: 8-core CPU (Intel i7 6th Gen or AMD Ryzen 5 2600).
- RAM: 16 GB or more (for multitasking and concurrent operations).
- Storage: 50 GB SSD (for fast data access and storage of embeddings).
- GPU: NVIDIA CUDA-enabled GPU with at least 6 GB VRAM (e.g., GTX 1660 or RTX 2060) for faster transcription and TTS synthesis.
- Audio Setup: High-quality microphone and speakers for clear voice capture and playback.
- Internet Connection: Recommended 20 Mbps or higher for seamless operation.

Chapter 4

Implementation

This chapter will detail the descriptions of the development of all iterations. The details will include all the necessary information pertaining to the development of a particular increment (Required Word Count: 2,500 to 3,500 Words). Such as:

1. Adopted Methodology for each iteration
2. Requirements of the iterations
3. Design, Code, and Testing results of each iteration
4. Dataset Details
5. Data Preprocessing
6. Data Analysis
7. Utilized Techniques
8. Evaluation Metrics, etc.

Chapter 5

Results and Discussions

This chapter covers the following (Required Word Count: 2,500 to 3,000 Words):

1. Achieved Results
2. Results Analysis
3. Results Comparisons
4. Expectations, etc.

Chapter 6

Conclusion

This chapter concludes the report (Required Word Count: 250 to 300 Words).

Appendix A

Extra Details

Supporting material of considerable length, lists, questionnaires, source code etc., which would interrupt the main text, should be included as appendix. Label appendices as A, B, etc.

Appendix B

Source Code

Add source code here.

References

[1] A. M. Turing, "Computing machinery and intelligence," *Mind*, vol. 59, no. 236, pp. 433–460, 1950. [Online]. Available: https://academic.oup.com/mind/article/LIX/236/433/986238

[2] M. McTear, *Conversational Interaction Systems: Design, Development, and Evaluation*. Springer, 2017.

[3] J. Weizenbaum, "Eliza - a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966. [Online]. Available: https://dl.acm.org/doi/10.1145/365153.365168

[4] K. Harrison, "Smart child: A chatbot for kids," *Computer Networks*, vol. 35, no. 3, pp. 291–304, 2001.

[5] G. A. Fowler, "Ibm's watson wins jeopardy," *The Washington Post*, 2011. [Online]. Available: https://www.ibm.com/history/watson-jeopardy

[6] S. Budge, N. Lutz, and S. Wexler, "Alexa, siri, cortana, and more: An introduction to voice assistants," *ResearchGate*, 2017.

[7] M. Rouse, "What is a chatbot?" 2020. [Online]. Available: https://www.techtarget.com/searchcustomerexperience/definition/chatbot

[8] Pakistan Institute of Education (PIE), "Pakistan education statistics 2021-22 report," 2022. [Online]. Available: https://www.pie.edu.pk/2021-22-report.pdf

[9] The Senate Standing Committee on Information Technology, "The senate standing committee on information technology," 2020. [Online]. Available: https://www.dawn.com/news/1569387

[10] "An overview of chat bot technology." Springer, 2020, pp. 373–383, first Online: 29 May 2020. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-49186-4$_3$1

[11] "Effects of generative chatbots in higher education," *Information*, vol. 14, no. 9, p. 492, 2023. [Online]. Available: https://www.mdpi.com/2078-2489/14/9/492

[12] M. R. Future, "Conversational artificial intelligence (ai) market overview," 2023, accessed on: 30 June 2023. [Online]. Available: https://www.marketresearchfuture.com/reports/conversational-ai-market-7913

[13] MarketsandMarkets, "Conversational (ai) market overview," 2023, accessed on: 30 June 2023. [Online]. Available: https://www.marketsandmarkets.com/Market-Reports/conversational-ai-market-49043506.html

[14] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong *et al.*, "A survey of large language models," *arXiv*, 2023.

[15] Y. Liu, T. Han, S. Ma, J. Zhang, Y. Yang, J. Tian, H. He, A. Li, M. He, Z. Liu *et al.*, "Summary of chatgpt/gpt-4 research and perspective towards the future of large language models," *arXiv*, 2023.

[16] A. Iyanda *et al.*, "A bilingual chatbot for home appliances control," in *AICTTRA 2023 Proceedings*, April 2023. [Online]. Available: https://www.researchgate.net/publication/379899921

[17] "Bilingual ai-driven chatbot for academic advising," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 13, no. 8, 2022. [Online]. Available: https://thesai.org/Publications/ViewPaper?Volume=13Issue=8Code=IJACSASerialNo=8

[18] X. Deng and Z. Yu, "A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education," *Sustainability*, vol. 15, no. 2940, 2023. [Online]. Available: https://www.mdpi.com/2071-1050/15/5/2940

[19] "Using a multilingual chatbot for foreign language practice," *Issues in Information Systems*, vol. 21, no. 2, pp. 85–92, 2020. [Online]. Available: https://pdfs.semanticscholar.org/0deb/3a1b0d187e4fe571c1e0ace9b99d6398c015.pdf

[20] N. J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*. Cambridge University Press, 2009.

[21] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2009.

[22] A. M. Turing, "On computable numbers, with an application to the entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. 42, no. 1, pp. 230–265, 1937.

[23] ——, "Computing machinery and intelligence," *Mind*, vol. 59, no. 236, pp. 433–460, 1950.

[24] J. McCarthy, M. Minsky, N. Rochester, and C. Shannon, "A proposal for the dartmouth summer research project on artificial intelligence," in *Dartmouth Conference on AI*, 1955.

[25] J. Weizenbaum, "Eliza—a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36–45, 1966.

[26] D. Crevier, *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books, 1993.

[27] E. A. Feigenbaum and P. McCorduck, *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*. Addison-Wesley, 1983.

[28] M. Campbell, A. J. Hoane, and F. hsiung Hsu, "Deep blue," *Artificial Intelligence*, vol. 134, no. 1-2, pp. 57–83, 2002.

[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[30] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[31] G. E. Hinton, A. Krizhevsky, and I. Sutskever, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[32] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface: Talking to Smart Devices*. Springer, 2016.

[33] "A literature survey of recent advances in chat bots," *Information*, vol. 13, no. 1, p. 41, 2023. [Online]. Available: https://www.mdpi.com/2078-2489/13/1/41

[34] G. Neff and P. Nagy, "Talking to bots: Symbiotic agency and the case of tay," *International Journal of Communication*, vol. 10, pp. 4915–4931, 2016.

[35] W. Chamberlain and J. Hall, *The Policeman's Beard is Half Constructed: Computer Prose and Poetry by RACTER*. New York: Warner Books, 1984.

[36] G. Güzeldere and S. Franchi, "Dialogues with colorful personalities of early ai," *Stanford Electronic Humanities Review*, vol. 4, no. 2, 1995.

[37] "Chatbots: History, technology, and applications," *ScienceDirect*, vol. 2, p. 106, December 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827020300062

[38] "Information on dr. sbaitso." [Online]. Available: https://classicreload.com/dr-sbaitso.html

[39] P. S. Al-Muhtadi and P. P. N. Gade, "Challenges in ai and nlp in chatbot applications," *Journal of Computer Science*, vol. 35, no. 3, pp. 215–222, 2018.

[40] A. M. Patel, "Security risks in ai-based chatbot applications," *Journal of Cyber Security and AI*, vol. 7, no. 1, pp. 22–33, 2022.

[41] L. J. Smith and R. M. Barrow, "Complexity in chatbot-based customer service systems," *Journal of Business Technology*, vol. 42, no. 2, pp. 134–145, 2019.

[42] S. J. Smith and A. Kumar, "Emotional intelligence in ai: The gap in current chatbot technology," *Journal of Human-Computer Interaction*, vol. 39, no. 4, pp. 78–89, 2022.

[43] H. Liu and T. P. N. Gera, "Human empathy versus chatbot empathy: A comparative study," *Journal of AI and Emotion*, vol. 12, pp. 115–123, 2021.

[44] K. R. S. R. Anitha and S. V. S. R. Srinivas, "Understanding chatbot technology," *International Journal of Computer Applications*, vol. 168, no. 6, pp. 6–10, 2017.

[45] R. W. Jones, "Data privacy concerns with chatbot applications," *Cybersecurity & Data Privacy*, vol. 18, pp. 40–52, 2020.

[46] M. H. Kumar and V. S. Verma, "Biases in machine learning algorithms: Effects on chatbot responses," *AI & Ethics Journal*, vol. 3, pp. 45–58, 2021.

[47] L. J. Smith, "Security concerns and privacy implications of chatbots," *Information Security Journal*, vol. 28, pp. 77–89, 2019.