

The Faculty of Informatics at Midocean University

AI Science Master's program.

**Customer support using Al technologies**

BY

**Maha Musaad Almuhaysin**

**Mouath Abdulrahman Alhanaki**

A Thesis Submitted to

Department of Informatics for the Master’s degree of AI Science

Under the Supervision of

**Prof. Dr. Yasser Abdelhamid**

Associate Professor, Faculty of Informatics, Midocean University

## Acknowledgements

We are grateful to Prof. Dr. Yasser Abdelhamid for his guidance and support throughout this research. His expertise has been instrumental in the successful completion of this study.

We also acknowledge the contributions of the Human Resources Development Fund (HRDF) in Saudi Arabia. Their commitment to innovative customer support solutions has enabled us to explore the integration of AI technologies within the Wusool program and uncover its transformative potential.

## Abstract

Providing efficient and effective customer support is a critical challenge for many organizations in today's highly competitive business landscape.

This project proposes the implementation of an advanced AI-powered system to significantly enhance the customer support experience by improving response times, automating routine inquiries, and increasing overall customer satisfaction. The initiative will be executed in five strategic phases: Requirement Gathering, Data Preparation and Model Development, System Integration, Deployment, and Evaluation.

Recent studies have shown that the integration of AI-based chatbots and virtual assistants in customer service can lead to substantial improvements in operational efficiency and customer experience. For example, “Between artificial intelligence and customer experience: a literature review on the intersection" which provides an overview of the current research at the intersection of AI and customer experience (CX) (Peruchini, M., da Silva, G.M. & Teixeira, J.M., 2024).

The review analyzes how technologies like chatbots, machine learning, and recommendation systems are being used to enhance the customer experience, particularly in sectors like tourism, banking, and e-commerce.

The proposed project aims to develop customer support system using the latest AI technologies to enhance the efficiency and effectiveness of the Wusool program offered by the Human Resources Development Fund (HRDF) in Saudi Arabia.

Key outcomes of this project include:

1. **Significant Reduction in Response Times**: The AI-powered customer support system will be able to handle routine inquiries and requests automatically, significantly reducing the response times for customers.
2. **Automation of Routine Inquiries**: The system will leverage natural language processing and machine learning algorithms to automate the handling of common customer inquiries related to the Wusool program, such as eligibility criteria, application process, and program benefits.
3. **Increase in Customer Satisfaction**: By providing efficient and effective customer support, the AI-powered system is expected to significantly improve customer satisfaction with the Wusool program. Customers will receive timely and accurate responses to their queries, enhancing their overall experience with the program.
4. **Enhanced Focus on Complex Interactions**: With the automation of routine inquiries, customer service agents will be able to dedicate more time and attention to handling complex and value-added interactions. This will enable them to provide more personalized support and address customers' unique needs and concerns.

By implementing this AI-powered customer support system for the Wusool program, the HRDF can significantly enhance the overall user experience, improve operational efficiency, and contribute to the broader goal of increasing employment opportunities and stability for their customers in the private sector and individuals with disabilities in Saudi Arabia.

## Table of Contents

[List of Abbreviations 10](#_Toc173004735)

[List of Figures 10](#_Toc173004736)

[List of Tables 10](#_Toc173004737)

**Chapter 1: Introcudion** ............................................................................... 9

1.1 Problem Definition ................................................................................... 11

1.2 Thesis Objectives .................................................................................... 14

1.3 Thesis Contribution ................................................................................ 15

1.4 Thesis Organization.................................................................................. 18

### Chapter 2: Background on CRM and AI Techniques ................... 20

2.1 Introduction ............................................................................................. 20

2.2 Application of AI in CRM systems........................................................... 22

2.2.1 Chatbots and Virtual Assistant......................................................... 23

2.2.2 Lead Scoring and Management......................................................... 24

2.2.3 Personalized Marketing.................................................................... 25

2.2.4 Process Optimization........................................................................ 25

2.2.5 Data Management............................................................................. 25

2.3 Chatbots vs. Virtual Assistants......................................................... 26

2.3.1 Case Studies (Chatbots vs. Virtual Assistants) ..............................26

2.3.2 IBM Watson Assistant................................................................27

2.3.3 Autodesk’s AVA .......................................................................32

**Chapter 3: Related Work**......................................................................... 37

**Chapter 4: Proposed Approach** ............................................................... 48

4.1 Step1: Data Acquisition and Cleaning................................................ 49

4.2 Step 2: Data Embedding and Storage.................................................50

4.3 Step 3: Question Retrieval and Contextual Response Generation......51

4.4 Technologies Used..............................................................................53

4.5 Implementation and Testing ...............................................................54

### Chapter 5: Experimental Results and Discussion ................................... 55

**Chapter 6:** **Conclusion and Future Work** ............................................. 59

6.1 Conclusion ...............................................................59

6.2 Future Work .............................................................60

References ................................................................................................ 64

# List of Abbreviations

* **AI:** Artificial Intelligence
* **CX:** Customer Experience
* **HRDF:** Human Resources Development Fund
* **Wusool:** Wusool program initiated by the Human Resources Development Fund (HRDF) in Saudi Arabia.
* **API:** Application Programming Interface
* **CRM:** Customer Relationship Management
* **ML:** Machine Learning
* **DL:** Deep Learning
* **AVA:** Autodesk Virtual Assistant
* **LLM:** Large Language Models

# List of Figures

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | **2.1** Comprehensive Architecture of a CRM System ………………………… | 22 | | **2.2.1** Architecture of a Chatbot System ……………………………………… | 23 | | **2.2.2** Machine Learning-Driven Personalized Advertising Workflow……… | 24 | | **3** Basic RAG flow ….…….….…….….….…….….….…….….….…….…. | 25 | | **3.1** Advanced Retrieval-Augmented Generation System with Knowledge Correction ….….…….….….…….….…….….…….….….…….….…….…. | 38 | | **3.2** LangGraph implementation for CRAG.….…….….……. .….…….…. | 39 | | **3.3** Timeline of early-stage (AI) software and models up to SOTA LLMs | 40 | | **3.4** The encoder-decoder structure of the transformer architecture (Ashish Vaswani | 41 | | **3.5** Three key steps in the LLM training process | 42 | | **3.6** Architecture of a Retrieval-Augmented Generation (RAG) system | 44 | |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Chapter 1 Introduction

Today's customer service landscape is rapidly evolving, with Artificial Intelligence (AI) at the heart of this transformation. As businesses expand, they often find themselves grappling with an increasing volume of customer inquiries. Traditional support systems like Customer Relationship Management (CRM) can falter under this pressure, leading to slower response times and a decline in the quality of service. Such shortcomings can significantly dampen customer satisfaction.

However, the integration of AI into the customer support systems is not merely a remedy for efficiency; it's a gateway to redefining customer interactions. By harnessing AI, organizations have the opportunity to not only streamline their processes but also to add a layer of personal touch to their communications, fundamentally enhancing the overall customer experience AI, particularly in the context of CRM, leverages machine learning (ML) and deep learning (DL) methods to extract insights from data, identify patterns, and make decisions with minimal human oversight (Kumar et al., 2020).

Also, the modern language generation model ChatGPT, created by Open Artificial Intelligence (AI), with the ability to “generate human-like text” (Aydın & Karaarslan, 2022) is recognized for its capacity to comprehend context and produce pertinent content. This model is built on the transformer architecture, which enables it to process massive volumes of data and produce text that is both cohesive and illuminating. Service is a crucial component everywhere as it provides the basis for establishing client rapport and offering aid and support.

Businesses may enhance customer experience by using ChatGPT's potential for assistance in any sector. The application of ChatGPT for customer support has been one of the most significant advances in recent years. This strategic adoption of AI technologies paves the way for more responsive, intuitive, and personalized service, transforming how businesses connect with their customers.

### Problem Definition

One of the persistent challenges in Customer Relationship Management (CRM) is providing personalized, timely responses to customer inquiries. Traditionally, CRM systems have been effective in managing customer data but often fall short in delivering real-time, context-aware interactions. This can lead to customer dissatisfaction, as modern consumers expect quick and tailored services based on their previous interactions and preferences.

Many organizations struggle with integrating their CRM systems across various departments. This often results in data silos where valuable customer information is trapped in one part of the business and inaccessible to others, hindering effective communication and service delivery (Ngelyaratan & Soediantono, 2022).

Scalability Issues, as businesses grow, the volume of customer interactions increases exponentially. Traditional CRM systems may not scale efficiently, leading to slower system performances and increased waiting times for customers.

As I want to focus here on the main problems faced in the Wusool program initiated by the Human Resources Development Fund (HRDF) of Saudi Arabia, aims to alleviate transportation costs for female employees in the private sector and male employees with disabilities. The program collaborates with licensed ride-hailing applications like Uber and Careem to facilitate the commuting needs of its beneficiaries. Despite its noble objectives and widespread adoption, the program faces several challenges related to customer support, which could be mitigated through the implementation of advanced AI technologies.

**Current Challenges:**

1. **High Volume of Inquiries**: The program receives a substantial number of inquiries regarding eligibility, registration, and technical issues.
2. **Response Delays**: Manual handling of queries often results in delayed responses, impacting user experience.
3. **Information Dissemination**: Users frequently encounter difficulties in accessing or understanding program details, eligibility criteria, and support mechanisms.
4. **Technical Issues**: Users face technical issues related to application integration, support activation, and maintaining account status.
5. **Resource Allocation**: Human resources are heavily taxed with repetitive tasks, reducing efficiency and increasing operational costs.

AI technologies, particularly those using machine learning algorithms, can analyze vast amounts of data from CRM systems to identify patterns and preferences of individual customers. ChatGPT can leverage this data to generate personalized responses automatically, ensuring that each customer feels understood and valued. This not only improves the customer experience but also enhances customer loyalty and retention (Libai et al., 2020)

Real-Time Response and Interaction, by Chat Generative Pre-Trained Transformer (ChatGPT) can process and respond to customer queries in real time, significantly reducing waiting times. With its advanced natural language processing capabilities, it can understand and engage in human-like conversations, providing immediate and relevant assistance to customers. This capability is crucial for maintaining high customer satisfaction levels in a fast-paced market environment (J. Paul, A. Ueno, C. Dennis, 2023)

ChatGPT can be integrated into existing CRM systems to bridge the gap between data silos within an organization. It can pull information from various sources to provide a comprehensive view of the customer, which is accessible to all relevant departments. This holistic approach ensures that all interactions with a customer are informed and consistent across all touchpoints (A.S. George, A.H. George, 2023).

AI-driven solutions like ChatGPT are highly scalable, capable of handling an increasing number of interactions without additional significant resources. This makes it an ideal solution for growing businesses that need to manage large volumes of customer interactions efficiently. Moreover, the automation of routine inquiries frees human agents to handle more complex issues, enhancing overall productivity and service quality. By addressing these core CRM issues effectively, AI and ChatGPT not only streamline customer relationship management processes but also transform them into more customer-centric systems that can adapt and evolve in response to changing customer needs (Abid Haleem, Mohd Javaid, Ravi Pratap Singh, 2024)

### Thesis Objectives

The objective is to develop and evaluate an advanced AI-powered customer support system that leverages state-of-the-art machine learning techniques to enhance user satisfaction and operational efficiency. Specifically, can contribute to one or more of the following objectives:

**Reduce Response Times:**

Implement AI-driven chatbots using deep learning and reinforcement learning to provide prompt and context-aware responses to customer inquiries.

**Automate Routine Inquiries:**

Utilize machine learning models such as BERT for the automatic classification and resolution of common customer support tickets, thereby minimizing the need for human intervention.

**Improve Sentiment Analysis:**

Apply advanced sentiment analysis models to extract nuanced insights from customer feedback, helping support teams better understand and address customer concerns.

**Enable Predictive Support:**

Integrate predictive analytics to anticipate and resolve potential customer issues before they escalate, allowing for proactive customer support.

**Enhance Multilingual Support:**

Incorporate AI translation models to facilitate seamless customer interactions in multiple languages, thereby improving the accessibility and inclusivity of the support system.

**Integrate Voice Assistants:**

Explore the use of AI-powered voice assistants to handle customer support interactions, making the support system more accessible and user-friendly.

**Detect Anomalies in Customer Behavior:**

Implement anomaly detection algorithms to identify unusual patterns in customer behavior or support interactions, enabling timely interventions and issue resolution.

By achieving these objectives, the proposed AI-powered customer support system aims to significantly improve the overall customer support experience, ensuring quicker, more accurate, and personalized responses to customer inquiries.

### 1.3 Thesis Contribution

The AI-powered customer support system will significantly address the challenges faced by the Wusool program. By leveraging advanced natural language processing (NLP) and machine learning (ML) algorithms, the system can handle a high volume of inquiries automatically, reducing the workload on human support agents and providing immediate feedback, thus enhancing the user experience. The system will efficiently disseminate information about the program, ensuring users have better access to details, eligibility criteria, and support mechanisms. It will also include diagnostic capabilities to address technical issues related to application integration, support activation, and account maintenance, guiding users through troubleshooting steps and escalating complex problems to human agents when necessary. Automating repetitive tasks will free up human resources to focus on more complex activities, improving overall efficiency and reducing operational costs.

So the Expected Outcomes and Benefits in general:

* Significant reduction in response times for customer inquiries
* Automation of routine inquiries related to Wusool program
* Increase in customer satisfaction with Wusool program
* Enhanced focus on complex and value-added customer interactions

This AI-powered solution is not limited to the Wusool program; it can be applied to any Customer Support system. By integrating AI technologies, organizations can handle increasing volumes of inquiries, reduce response times and operational costs, provide personalized responses, ensure uniform information dissemination, and derive insights from user interactions to inform continuous improvement and strategic decision-making.

### 1.4 Thesis Organization

This thesis is organized into six chapters, each focusing on different aspects of AI-powered customer support systems and their application within CRM environments, specifically within the context of the Wusool program.

**Chapter 1: Introduction**

This chapter sets the stage for the thesis, providing a comprehensive introduction to the research topic. It includes the problem definition, thesis objectives, contributions, and an overview of the thesis organization.

* **1.1 Problem Definition:** This section outlines the specific challenges and issues that the thesis aims to address. It details the limitations of current customer support systems and the need for AI integration.
* **1.2 Thesis Objectives:** This part specifies the primary goals and aims of the research. It highlights the intended outcomes and what the research seeks to achieve.
* **1.3 Thesis Contribution:** This section describes the unique contributions of the thesis to the field of AI-powered customer support. It emphasizes the novel aspects of the research and its impact on CRM systems.
* **1.4 Thesis Organization:** Provides an overview of the structure of the thesis, summarizing the content of each chapter.

**Chapter 2: Background on CRM system and AI Techniques**

This chapter provides an overview of AI applications in CRM systems, emphasizing the various aspects where AI technology enhances customer relationship management.

* **2.1 Introduction:** A general introduction to AI-powered CRM systems.
* **2.2 Application of AI in CRM systems:** Detailed examination of different AI applications within CRM systems.
  + **2.2.1 Chatbots and Virtual Assistants:** Explores the role and functionality of chatbots and virtual assistants in automating customer interactions.
  + **2.2.2 Lead Scoring and Management:** Discusses how AI can be used to score and manage leads more effectively.
  + **2.2.3 Personalized Marketing:** Examines the use of AI in creating personalized marketing strategies.
  + **2.2.4 Process Optimization:** Looks at how AI optimizes various processes within CRM systems.
  + **2.2.5 Data Management:** Explores the role of AI in managing and analyzing customer data.
* **2.3 Case Studies on (Chatbots vs. Virtual Assistants):**

Comparison between chatbots and virtual assistants in terms of functionality and use cases

* + **2.3.1 IBM Watson Assistant:** Detailed case study on IBM Watson Assistant.
  + **2.3.2 Autodesk’s AVA:** Case study on Autodesk’s AVA and its implementation.

**Chapter 3: Related Work**

This chapter delves into the existing article and research related to the use of embedding techniques and Large Language Models (LLMs) in various applications. It critically evaluates how these technologies have been integrated into systems and identifies areas where further research could advance the field. The discussion situates this thesis within the wider scholarly framework, highlighting its contribution to the understanding and application of embeddings and LLMs in new and existing domains.

The research paper also explores the use of prompt engineering and reinforcement learning techniques to enhance the performance and capabilities of these language models in specific applications.

**Chapter 4: Proposed Approach**

This chapter describes the proposed approach used in the thesis. It outlines the design and development of the AI-powered chatbot for the Wusool CRM, including data extraction, vectorization, retrieval systems, and integration with OpenAI's GPT and embeddings.

**Chapter 5: Results and Discussion**

This chapter presents the findings of the research. It includes a detailed analysis of the data collected and discusses the implications of the results in the context of AI-powered customer support systems.

**Chapter 6: Conclusion and Future Work**

The final chapter summarizes the key findings of the thesis, discusses the contributions, and suggests areas for future research. It reflects on the limitations of the study and provides recommendations for further investigation.

## Chapter 2 Background on CRM system and AI Techniques.

### 2.1 Introduction

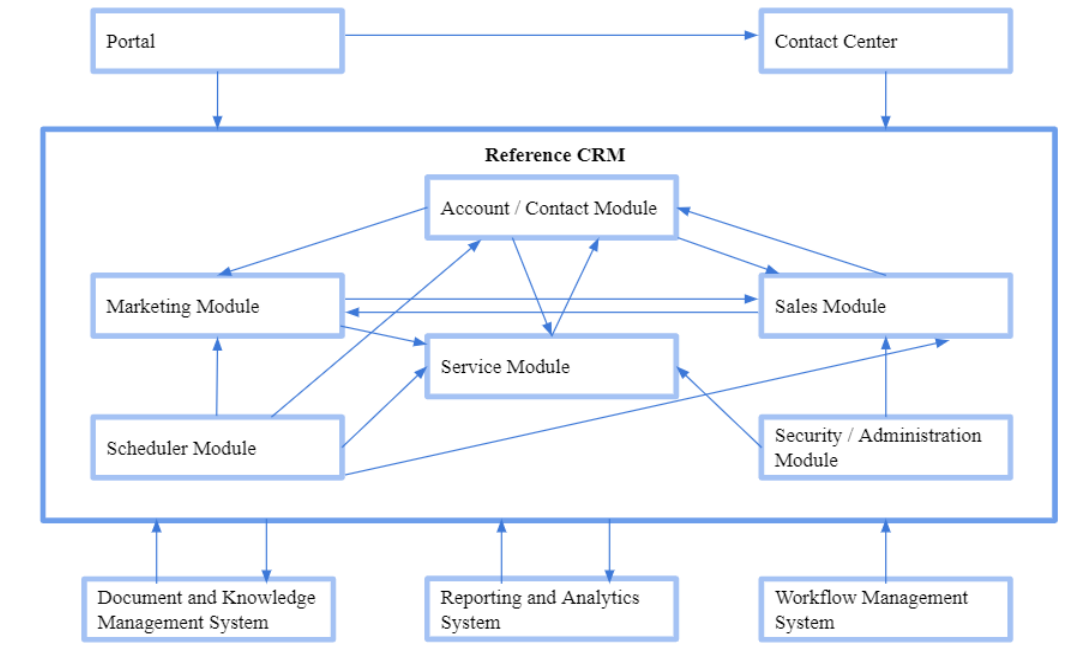
Customer Relationship Management (CRM) is a strategic approach that focuses on collecting, managing, and intelligently utilizing data with the support of technology solutions to develop valuable relationships with key customers and provide an exceptional customer experience.

The next significant development toward a unique and more effective CRM is the integration of artificial intelligence (AI) (Kumar et al., 2020). AI, particularly in the context of CRM, involves leveraging machine learning (ML) and deep learning (DL) techniques to extract insights from data, recognize patterns, and make decisions with limited human intervention (Kumar et al., 2020), or even, in the case of DL, to learn from mistakes without human involvement (Zaki, 2019). Successful companies have effectively employed AI in CRM for various applications, such as customer data analysis, chatbots and virtual assistants, personalized recommendations, sentiment analysis, computer vision, voice and speech recognition, and predictive analytics (Libai et al., 2020; Abousaber, 2023; Kumar et al., 2020). According to a recent article published in the Harvard Business Review, AI technologies can significantly impact companies' key areas, including forecasting, performance management, upselling, and cross-selling (Antonio, 2018).

**Figure 2.1.1** below depicts a CRM application architecture for, proposed by Cruz & Vasconcelos (2015). The architecture consists of various interconnected modules and systems that collectively support the functionality of the CRM. Beyond immediate benefits, the integration of AI in CRM also holds long-term strategic implications, enabling companies to adapt and thrive in an ever-changing market landscape. Understanding and addressing the challenges associated with this integration are essential for businesses to make informed decisions and prevent technology investments from being made without adequate awareness of the organizational and managerial changes required. Currently, the organizational and managerial elements that should be considered in the design and execution of these cutting-edge systems within the CRM context are not completely understood by practitioners. Uncertain outcomes, high levels of perceived complexity, and lack of experience are some reasons why AI integration in CRM is still low (Mishra et al., 2022). Executives struggle not only with the complexity of the technology but also with the organizational and managerial challenges of integrating AI into established organizational routines.

To find potential use cases for AI technologies, it is important to understand

these modules forming a CRM system and the ones interacting with it.



**Figure 2.1.1** Comprehensive Architecture of a CRM System (Cruz and Vasconcelos, 2015)

In the end, the CRM system has a significant role on both enterprise level and the various business functions of the company.

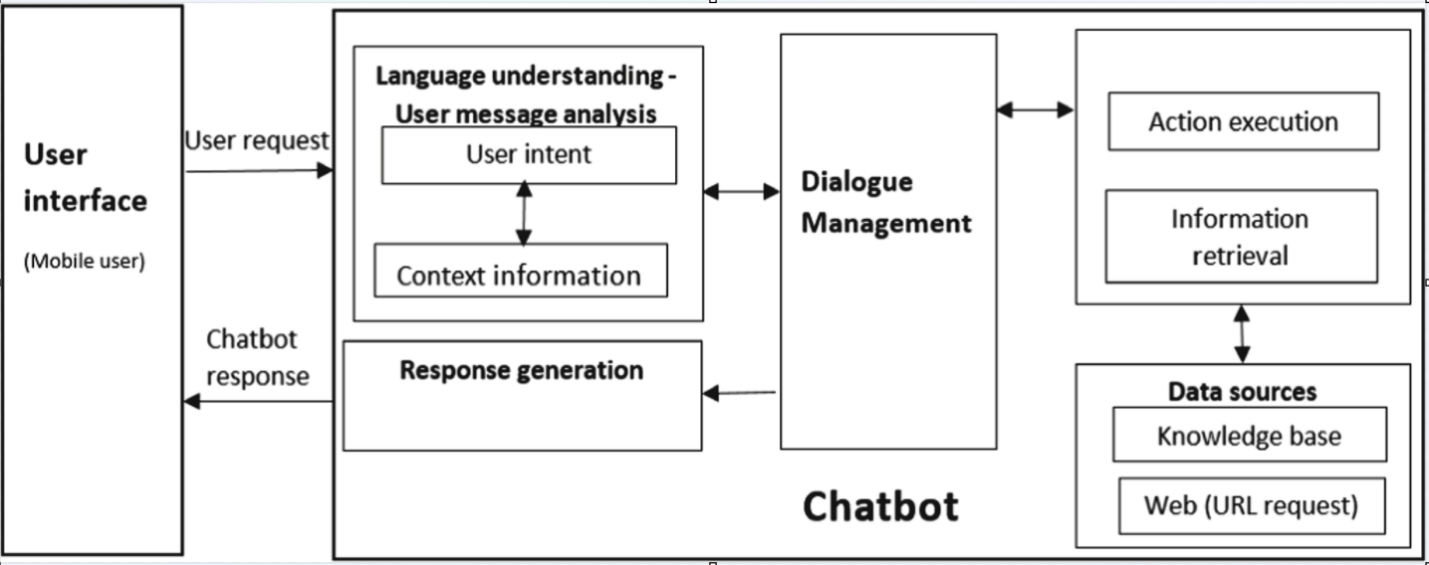
### 2.2 Applications of AI in CRM systems

Artificial Intelligence (AI) is transforming the way customer relationship management (CRM) systems work. By integrating AI capabilities, CRMs can automate and enhance various customer support functions, leading to improved efficiency and better customer experiences.

Some key applications of AI in CRM include:

#### 2.2.1 Chatbots and Virtual Assistants

AI-powered chatbots and virtual assistants can provide 24/7 customer support, answering common queries and guiding customers through processes like troubleshooting or making purchases. These AI agents use natural language processing to understand customer requests and provide relevant responses. (S V Sai Abitha, 2021)

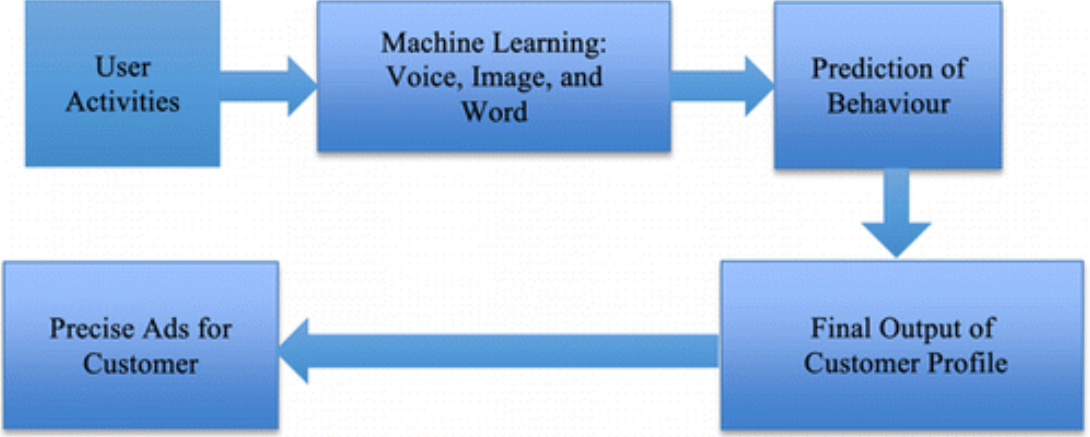
In figure 2.2.1 below the architecture of a chatbot system. It describes the flow of interaction between a user (specifically a mobile user) and the chatbot by highlighting the key components and processes involved in understanding and generating responses to user requests.

**Figure 2.2.1** (Architecture of a Chatbot System by S V Sai Abitha, 2021)

#### 2.2.2 Predictive Analytics

AI can analyze customer data, such as purchase history and behavior, to predict future customer needs and preferences. This allows businesses to proactively engage with customers and offer personalized solutions, improving customer satisfaction and loyalty. (Siti Zulaikha and Hazik, 2020)

By AI we can figure out the content and conclude the results based on the keywords used, semantic index, and synonyms as shown in Figure 2.2.2. In addition, AI can automatically understand the intensity of users to track their behaviors and future predictions.



**Figure 2.2.2** Machine Learning-Driven Personalized Advertising Workflow (Siti Zulaikha and Hazik, 2020)

#### 2.2.3 Lead Scoring and Management

AI can help CRMs automatically score and prioritize leads based on their likelihood to convert, allowing sales teams to focus on the most promising prospects and improve overall sales efficiency. (Kapil Kumar Sharma, 2023)

#### 2.2.4 Personalized Marketing

AI can segment customers based on their data and behavior, enabling businesses to deliver highly targeted and personalized marketing campaigns, which can increase engagement and conversion rates. ([Gao, Y.](https://www.emerald.com/insight/search?q=Youjiang%20Gao) and [,](https://www.emerald.com/insight/search?q=Hongfei%20Liu) 2023)

#### 2.2.5 Process Optimization

AI can help identify inefficiencies in CRM workflows and suggest improvements, leading to streamlined operations and more timely, relevant customer interactions. (Timone Silviu STĂNCIOIU, 2023)

#### 2.2.6 Data Management

AI can automate data entry, cleansing, and enrichment processes, ensuring the accuracy and consistency of customer data within the CRM system, which is crucial for effective personalization and analytics. (Fangyuan Li, 2022)

By leveraging these AI-powered capabilities, businesses can enhance their customer support, sales, and marketing efforts, ultimately improving the overall customer experience and driving business growth. (Keegan, B.J., Canhoto, A.I., Yen, D.A. wan, 2022)

#### 2.3 Case Studies (Chatbots vs. Virtual Assistants)

AI technologies, particularly chatbots and virtual assistants, have gained prominence in customer service due to their ability to provide instant responses, automate routine tasks, and deliver personalized interactions. These technologies leverage Natural Language Processing (NLP), machine learning algorithms, and speech recognition to understand and respond to customer inquiries effectively

Chatbots are primarily text-based conversational agents that are designed to handle specific, routine inquiries and tasks (Luo, B., Lau, R. Y., Li, C. P., & Si, Y. W, 2022).

They are often used for customer service, providing answers to frequently asked questions, and assisting with simple transactions (Bavaresco, R., Silveira, D., Reis, E., Barbosa, J., Righi, R., Costa, C., ... & Vanzin, M. 2020).

Chatbots typically have a limited domain of knowledge and operate based on predefined rules or scripts, making them well-suited for repetitive, high-volume interactions (Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. 2020).

In contrast, virtual assistants are more advanced, AI-powered agents that can engage in both text-based and voice-based interactions. They have broader capabilities, allowing them to perform a wide range of tasks such as scheduling appointments, setting reminders, controlling smart home devices, and even providing personalized recommendations. Virtual assistants leverage natural language processing and machine learning algorithms to understand user intent, learn from interactions, and adapt their responses accordingly.

The key differences between chatbots and virtual assistants can be summarized as follows:

1. Interaction modality: Chatbots are primarily text-based, while virtual assistants can handle both text and voice interactions.
2. Complexity of tasks: Chatbots are typically limited to handling routine, predefined tasks, while virtual assistants can perform more complex, open-ended tasks.
3. Adaptability and learning: Virtual assistants have more advanced AI capabilities, allowing them to learn and adapt to user preferences over time, whereas chatbots often have a more rigid, rule-based approach.
4. Integration and platform: Chatbots are often integrated into specific applications or websites, while virtual assistants are typically integrated into broader platforms, such as smart speakers or mobile operating systems.

Chatbot integrated into a CRM system can help sales representatives by automatically updating customer records with the latest interactions and information. This ensures that the sales team is always up to date and can provide more tailored and informed responses to customers.

Virtual assistants can also help with more complex tasks, such as managing inventory, predicting demand, and optimizing supply chains. They can analyze data, identify trends, and provide recommendations to improve efficiency and reduce costs.

**To better understand the impact of AI-powered chatbots and virtual assistants in CRM, let’s look at a couple of real-world examples:**

### 2.3.1 IBM Watson Assistant

IBM Watson Assistant is an AI-powered virtual assistant used in various industries, including finance and healthcare. It’s capable of answering complex questions and can be integrated into CRM systems to provide personalized support to customers. For instance, in the financial sector, Watson Assistant can help customers with account-related inquiries, investment recommendations, and even fraud detection.

IBM Watson Assistant can be integrated with CRM systems to provide personalized support in the financial sector:

* Account-related inquiries:

Watson Assistant can be trained to understand and respond to a wide range of account-related questions, such as account balances, transaction histories, and payment due dates.

By integrating with the CRM system, Watson Assistant can access the customer's account information and provide accurate, real-time responses to their inquiries.

This helps to improve customer satisfaction and reduce the burden on human customer service representatives.

* Investment recommendations:

Watson Assistant can leverage the customer's financial data stored in the CRM system, such as risk tolerance, investment goals, and portfolio holdings.

Using this information, Watson Assistant can provide personalized investment recommendations and advice, helping customers make informed financial decisions.

This can include suggestions for portfolio rebalancing, new investment opportunities, or adjustments to the customer's investment strategy.

* Fraud detection:

By integrating with the CRM system, Watson Assistant can monitor the customer's account activity and transaction patterns.

Using advanced analytics and machine learning algorithms, Watson Assistant can detect potentially fraudulent activities, such as unauthorized transactions or suspicious spending behavior.

When a potential fraud incident is identified, Watson Assistant can immediately alert the customer and provide guidance on next steps, helping to mitigate the impact of fraud and protect the customer's financial well-being.

The seamless integration of Watson Assistant with CRM systems in the financial sector allows for a more personalized and proactive customer experience. Customers can receive timely and accurate responses to their inquiries, tailored investment recommendations, and enhanced fraud protection - all through a conversational AI interface that is available 24/7. This can lead to increased customer satisfaction, loyalty, and trust in the financial institution.

 Here's a more detailed breakdown of how it works, its components, and the various ways it can be integrated:

**How IBM Watson Assistant Works**

IBM Watson Assistant is built on a foundation of natural language processing (NLP) and machine learning (ML), which enables it to understand, interpret, and respond to human language. Here’s how it operates within a CRM system:

1. **Data Ingestion**: Watson Assistant ingests and processes data from various sources, including direct user inputs (through chat interfaces) and backend databases (like those maintained by CRM systems).
2. **Understanding Context**: Using NLP, Watson Assistant analyzes the text provided by the user to grasp the context and intent behind the inquiries or statements.
3. **Processing and Decision-Making**: Based on the understood intent and the data available from the CRM system, Watson Assistant applies predefined rules and ML models to generate responses, recommendations, or alerts.
4. **Response Generation**: Finally, the assistant formulates a response in natural language. This response is then delivered to the user through the interface they are interacting with, such as a web chatbot or mobile app.

**Components of IBM Watson Assistant**

1. **Intent Recognition**: Identifies the purpose of the user’s input using NLP techniques. This helps in understanding what the user is asking or stating.
2. **Entity Recognition**: Detects and categorizes key pieces of information in the user's input, such as dates, amounts, or specific financial terms.
3. **Dialogue Management**: Handles the flow of conversation, ensuring that the responses are coherent and contextually appropriate across multiple exchanges.
4. **Integration Layer**: Connects Watson Assistant to various data sources, including CRM databases, to retrieve and update customer-specific information.
5. **Analytics**: Monitors interactions to provide insights into usage patterns, common inquiries, and system performance, aiding continuous improvement.

**Integration Methods**

The integration of Watson Assistant with CRM systems can occur through several methods, each serving different integration depths and purposes:

1. **API Integration**: Watson Assistant can communicate with CRM systems via application programming interfaces (APIs). This allows real-time data exchange, where Watson Assistant can query the CRM database for customer data, update records, or trigger actions within the CRM.
2. **Middleware or Integration Platforms**: Using tools like IBM App Connect or other middleware solutions, Watson Assistant can be linked with CRM systems. This method is useful for more complex scenarios involving multiple systems (e.g., CRM, ERP, and external databases).
3. **Webhooks**: Watson Assistant can use webhooks to send and receive data from the CRM system. This is particularly useful for triggering real-time actions based on conversation flow, such as sending alerts or initiating workflows.
4. **Direct Database Access**: In some configurations, Watson Assistant may have direct access to the CRM database to read or write data. This requires careful security and privacy considerations but can reduce response times and complexity.

**Capabilities and Advantages**

* **Personalization**: By accessing comprehensive customer profiles in the CRM, Watson Assistant can tailor conversations and recommendations, making interactions more relevant and personal.
* **Scalability**: AI capabilities allow handling of a large volume of queries simultaneously, which can significantly reduce the load on human agents.
* **24/7 Availability**: Unlike human agents, Watson Assistant can provide continuous service, ensuring that customer inquiries are addressed at any time, improving overall customer satisfaction.
* **Fraud Detection**: Advanced analytics enable Watson Assistant to identify anomalies and potential fraudulent activities by analyzing transaction patterns and historical data.

The integration of IBM Watson Assistant with CRM systems enhances the efficiency and effectiveness of customer service operations in the financial sector, making it a powerful tool for improving customer engagement and security. (https://cloud.ibm.com/docs)

### 2.3.2 Autodesk’s AVA

Autodesk, a leader in 3D design, engineering, and entertainment software, developed a chatbot called AVA (Autodesk Virtual Assistant).

AVA assists customers with product inquiries, troubleshooting, and license management. This chatbot has significantly improved customer support efficiency, reduced response times, and provided a more consistent experience for customers.

Some of the key benefits that AVA has delivered for Autodesk include:

* Reduced response times: With AVA, Autodesk has been able to reduce the average time-to-resolution from one and a half days down to just minutes.

This significant improvement in efficiency has been a major advantage for customers.

* Improved customer satisfaction: As a result of the faster response times, Autodesk has seen a 10-point increase in customer service satisfaction levels. Customers appreciate the quick and reliable support they receive from AVA.
* High volume of interactions: AVA handles over 35,000 customer questions every month, making it the highest volume point of contact for Autodesk customers.

This demonstrates the chatbot's ability to scale and handle a large number of inquiries.

* Consistent experience: By using AVA, Autodesk is able to provide a more consistent experience for customers, as the chatbot delivers accurate, on-brand responses based on its extensive knowledge base.

Autodesk is further enhancing AVA's capabilities by integrating it with Soul Machines, a company that specializes in creating digital human avatars. This will give AVA a more human-like face, voice, and ability to recognize and respond to emotional cues, providing an even more natural and personalized customer experience.

Overall, the implementation of AVA has been a significant success for Autodesk, demonstrating the transformative impact that AI-powered chatbots can have on customer support efficiency, response times, and the overall customer experience.

Powered by natural language processing and machine learning, AVA’s knowledge base is comprised of chat logs, use cases and forum posts, thousands of keywords, phrases, clusters, as well as syntax and idioms. All of this data helps AVA understand the subtleties and nuances of language and what is being asked of it.

And with experience and fine-tuning, AVA gets smarter and faster over time.

Here’s a detailed breakdown of how AVA works, its components, and various integration methods:

**How AVA Works**

1. **Data Ingestion**: AVA starts by ingesting data from user inputs and Autodesk’s databases. This includes user queries, product details, user manuals, FAQs, and support documents.
2. **Language Processing**: Using NLP, AVA analyzes the text provided by the user. It identifies the intent of the user's query and extracts relevant entities (specific terms or phrases related to Autodesk products).
3. **Decision Making**: Based on the identified intent and the information extracted, AVA uses pre-set rules and machine learning models to determine the most appropriate response or action. This could involve answering a question, directing the user to a resource, or initiating a troubleshooting protocol.
4. **Response Generation**: AVA constructs a response in natural language, which is displayed to the user. This response could also include visual aids like diagrams or links to online resources if the chat interface supports those functions.

**Components of AVA**

1. **Intent Recognition**: This component identifies what the user is trying to achieve with their query, such as seeking help with a software error or inquiring about license details.
2. **Entity Recognition**: This involves identifying and classifying key pieces of information in the user's input, such as product names, error codes, or specific features.
3. **Dialogue Management**: Manages the flow of the conversation to ensure that responses are contextually appropriate, and that the conversation progresses logically.
4. **Integration Layer**: Facilitates communication between AVA and Autodesk’s internal systems, databases, and possibly third-party applications.
5. **Analytics**: Provides insights into the types of queries being handled, user satisfaction, resolution rates, and other key performance indicators.

**Integration Methods**

1. **API Integration**: AVA can be integrated with other systems via APIs, allowing it to pull information from Autodesk databases or initiate actions in other software tools used within the company.
2. **Middleware Solutions**: AVA might connect to other systems using middleware platforms that facilitate data synchronization and workflow automation across different software applications.
3. **Webhooks**: These allow AVA to send real-time data to other systems whenever specific events occur within the chatbot, such as the completion of a troubleshooting interaction.
4. **Direct Database Access**: In some cases, AVA may need to directly query a database to retrieve the most current data or to update records based on the conversation outcomes.

**Practical Applications within Autodesk**

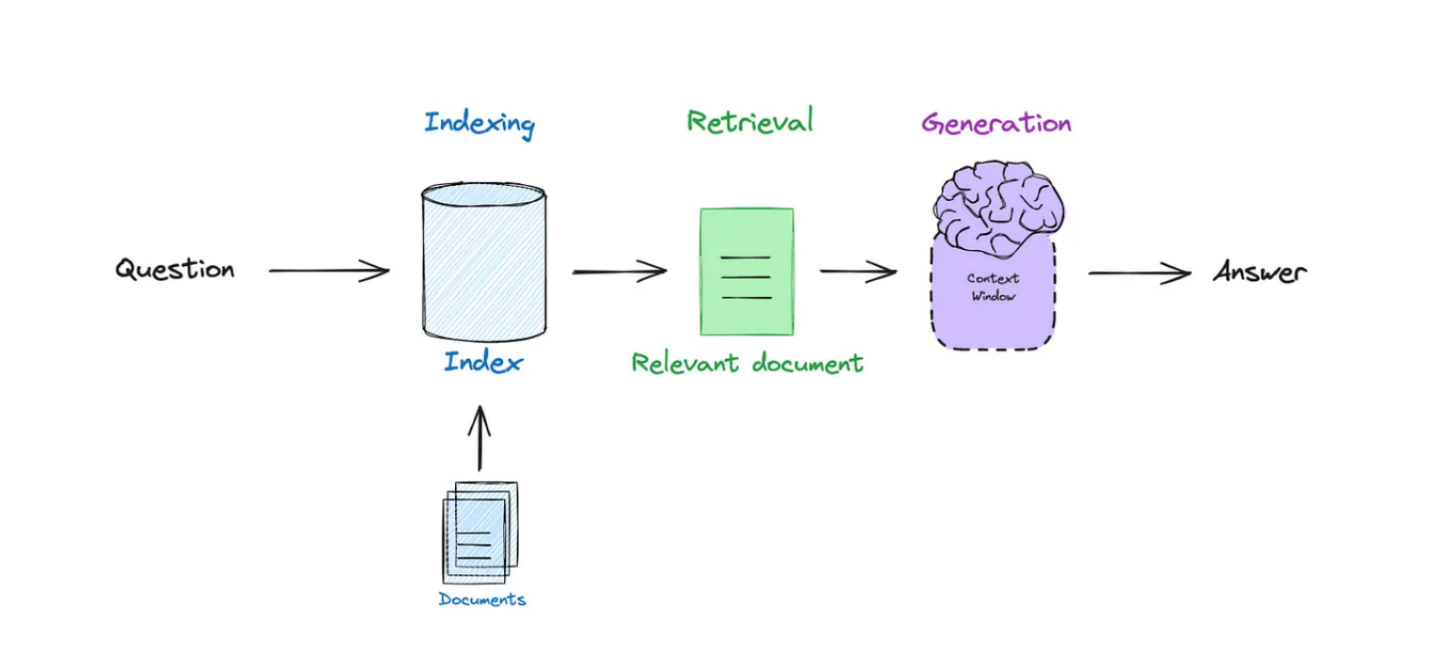
* **Customer Support**: Directly assisting users by providing answers to common questions, troubleshooting steps for software issues, or detailed product information.
* **Sales and Marketing**: Guiding potential customers through product options, pricing structures, and purchase processes.
* **Feedback Collection**: Gathering user feedback on Autodesk products and services, which can be used to inform product development and customer service strategies.
* **Training and Onboarding**: Helping new users learn how to use Autodesk products through interactive, conversational tutorials and tips.

(https://www.autodesk.com/)

## Chapter 3 Related Work

The article by (Cole McIntosh, 2024) discusses advancements in AI technologies, particularly focusing on enhanced versions of the Retrieval Augmented Generation (RAG) model as well as how we use it in this project.

* **Traditional RAG**: This model enhances AI-generated text by integrating language model outputs with information retrieved from a relevant document corpus. It aims to produce more contextually aware and content-rich outputs but struggles with the precision of retrieved documents, impacting the quality of responses.



**Figure 3** Basic RAG flow (<https://blog.langchain.dev/agentic-rag-with-langgraph/> , 2024)

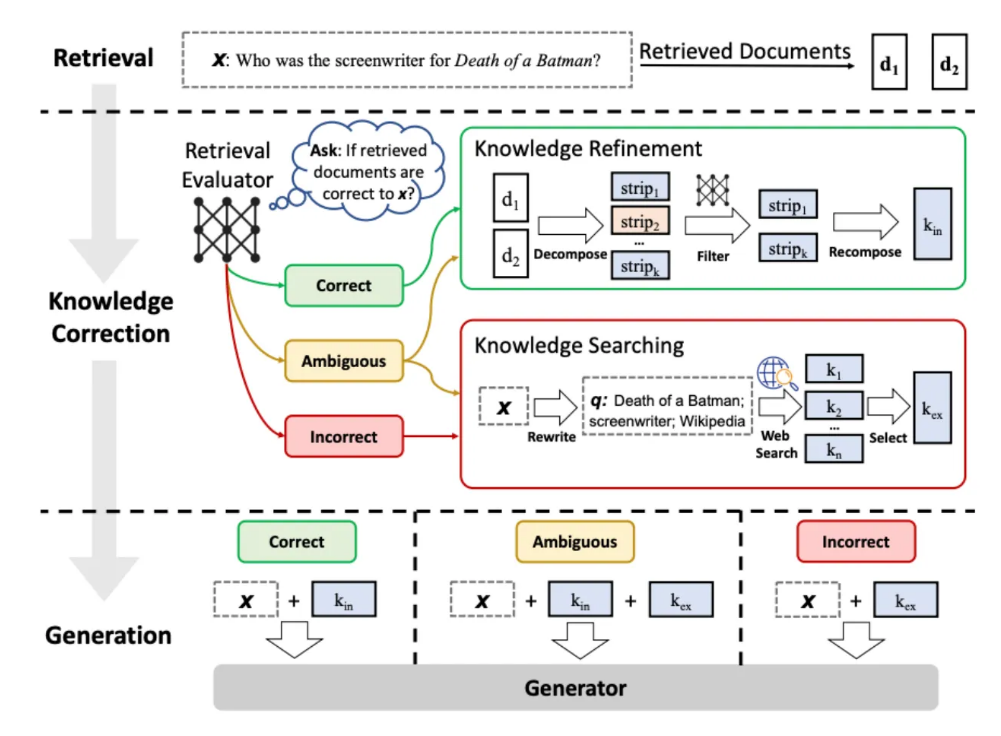
in **Figure 3** depicts a simplified flowchart illustrating the process of Retrieval-Augmented Generation (RAG) for answering a question. Here's a breakdown of the components and flow:

1. **Question**: This is the starting point where a user question is presented.
2. **Indexing**: The flowchart shows a cylinder labeled "Index," indicating a database or repository where documents are indexed for quick retrieval. This step involves the organization of documents to facilitate efficient searching.
3. **Retrieval**: This step involves selecting a "Relevant document" from the index based on its relevance to the question asked. The document is represented as a green rectangle.
4. **Generation**: The final step involves a brain-like icon labeled "Content Window," suggesting the use of the content from the relevant document to generate an answer. The process culminates in producing an "Answer" to the original question.

This is the basic workflow of a Retrieval-Augmented Generation system, demonstrating how a question is processed through indexing, retrieval of relevant documents, and finally, the generation of an answer based on the contents of the retrieved document. This visual aids in understanding the sequential steps involved in using indexed information to respond to queries dynamically.

**The Advanced RAG Methodologies:**

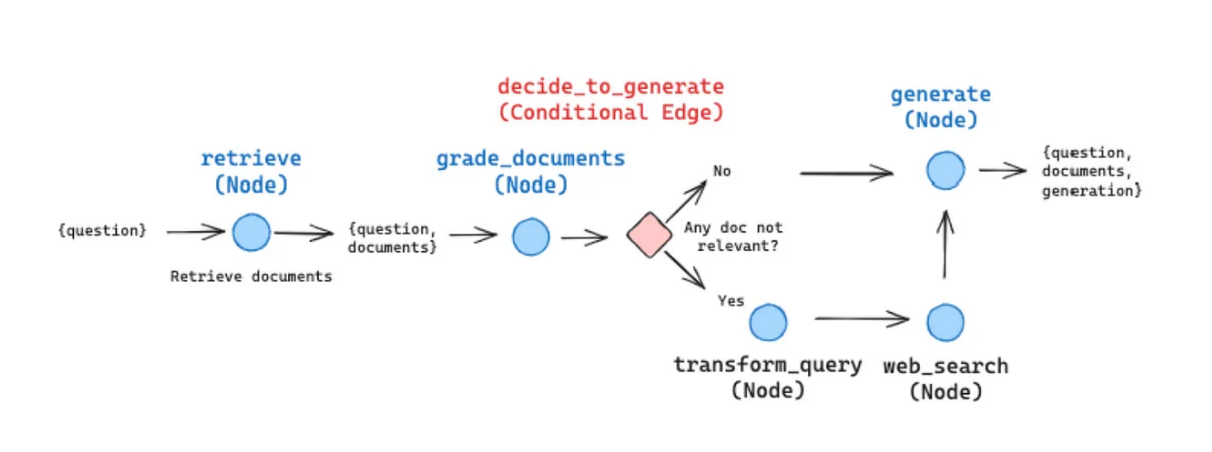
* **Corrective RAG (CRAG)**: This advanced form introduces a "corrective" mechanism that evaluates the relevance of retrieved documents before their use in generating responses. This ensures the accuracy and contextual relevance of the AI's output. CRAG also includes web searches as a fallback, enhancing the dynamism of the system in sourcing up-to-date information.
* **Self-Reflective RAG**: It integrates feedback loops into the RAG process, allowing the system to re-query or re-retrieve documents based on the initial output's relevance and accuracy. This method uses state machines to define a series of steps and transitions to improve content relevance and quality.

And as shown in **figure 3.1** the diagram illustrates the sophisticated process of a corrective Retrieval-Augmented Generation system, focusing particularly on handling ambiguities and inaccuracies in retrieved documents. It demonstrates how a system can evaluate, refine, and supplement information to ensure the generation of accurate and relevant answers. This model is particularly useful for complex queries where initial document retrieval may not provide sufficient or accurate information directly.

**Figure 3.1** Advanced Retrieval-Augmented Generation System with Knowledge Correction (<https://blog.langchain.dev/agentic-rag-with-langgraph/> , 2024)

**Role of LangGraph**

* **LangGraph**: This tool is vital in implementing CRAG and Self-Reflective RAG. It uses a graph-based approach to visualize and manage workflows, defining states, nodes, and edges that help in creating flexible and controlled AI systems. LangGraph simplifies the conceptualization and implementation of these methodologies.



**Figure 3.2** LangGraph implementation for CRAG (<https://blog.langchain.dev/agentic-rag-with-langgraph/>, 2024)

Also, there are various studies highlighting the benefits of using LLM chatbots like the research provided by (Debadutta Dash, Rahul Thapa, 2023).

He explained the LLM-based chatbots in customer support, LLM-powered chatbots can provide customers with instant responses to queries and can help resolve issues more quickly and efficiently than traditional customer service chatbots.

The research mentioned the main models and methodologies that are relevant to explore, personalize and analyze question-answering LLM-based chatbots:

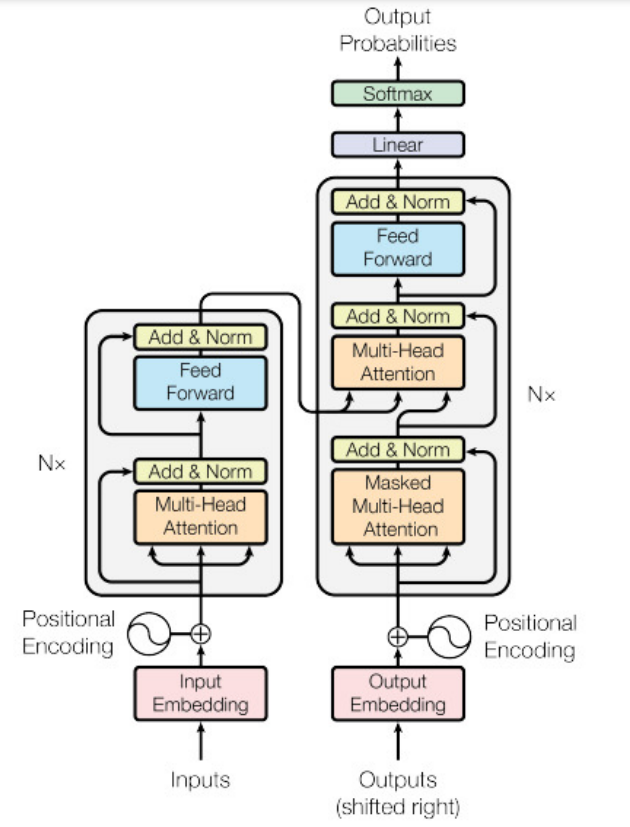
**Large Language Models (LLMs)** are a subset of AI models specifically designed to understand and generate human-like text based on vast amounts of data. These models, typically built using deep learning techniques, contain an extensive number of parameters, often in the range of billions or even trillions. The history of LLMs starts well before they became widely recognized by the general public during the release of ChatGPT in November 2022, as seen in **figure 3.3**

A colorful timeline with pointers

Description automatically generated

**Figure 3.3** Timeline of early-stage (AI) software and models up to SOTA LLMs (Rohan Anand, 2023)

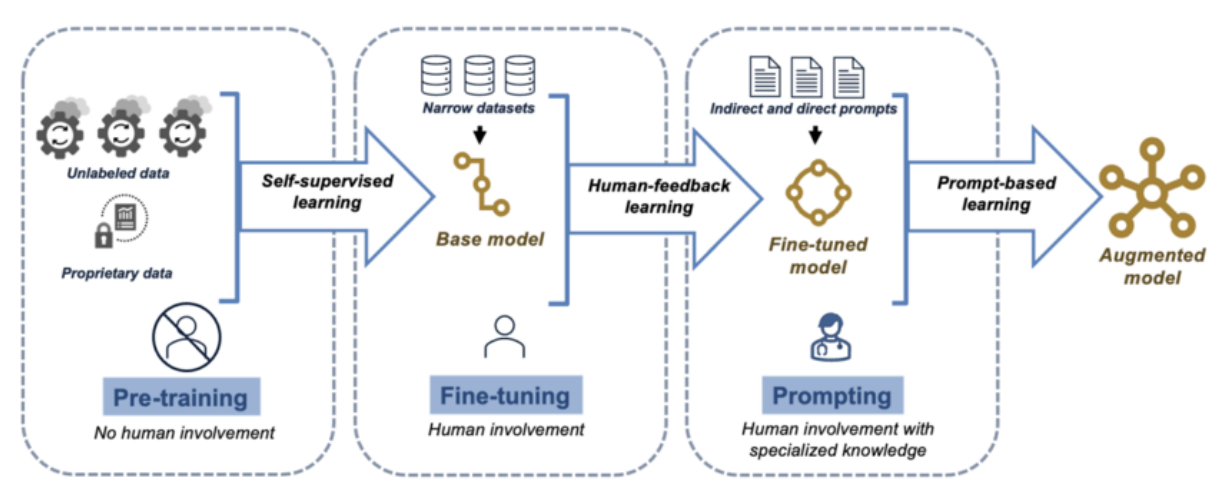
The heart of LLMs is the transformer architecture that includes a self-attention mechanism. **Figure 3.4** illustrates the general architecture of the transformer, which is composed of an encoder and a decoder. The encoder takes the input sequence and converts it into a series of continuous representations, while the decoder generates the output sequence from these representations. The key elements of this architecture will be further explained.



**Figure 3.4** The encoder-decoder structure of the transformer architecture (Ashish Vaswani, 2020)

**Training of LLMs** involves a series of steps and techniques that might be combined in different variations to meet the user’s specific needs. The training process, as shown in **figure 3.5**, can generally be divided into three key parts:

pre-training, fine-tuning and prompting. LLMs undergo an initial training phase called pre-training where the model is trained on a massive (unlabelled) dataset, often consisting of large parts of the internet. The LLM builds a foundational understanding of how language works, as it "learns" to recognize patterns and structures in the input data. After pre-training, models can be fine-tuned on specific tasks using more narrow (task-specific) datasets. The aim here is to have the model specialize in tasks like translation, question-answering or chatbot functionalities. Prompting refers to the process of interacting with a fine-tuned model using specific queries or statements (prompts) that guide the model to produce a desired output. These three key elements of the LLM training process and the techniques needed to train LLMs as question-answering chatbots will be further elaborated on in the subsequent sub-chapters.



**Figure 3.5** Three key steps in the LLM training process (Nature, 2023)

In their study, researchers (Debadutta Dash and Rahul Thapa, 2023) also explored combining dataset fine-tuning with another method known as Retrieval Augmented Generation (RAG), illustrated in **figure 3.6**. This approach involves integrating a chatbot with a database or knowledge base that is rich with the latest company-specific information, including details about products, services, policies, frequently asked questions, and even user contributions from community forums. When a customer asks a question, the chatbot employs RAG to draw pertinent information from this extensive database, weaving it into the response.

This ensures that the answers provided are not only relevant to the context but are also updated with the most current information, which is particularly useful for addressing queries related to recent changes, specific product details, or complex issues that demand in-depth knowledge.

In practical applications, a company may opt to simultaneously utilize both fine-tuning and RAG.

While fine-tuning helps the chatbot to navigate and respond suitably within the specific scenarios of the company's customer service, RAG enhances its capability by sourcing real-time, detailed content from external databases.

A diagram of a process

Description automatically generated

**Figure 3.6** Architecture of a Retrieval-Augmented Generation (RAG) system (Epsilla x LangChain, 2023)

Also, one of the important techniques used in this project is the prompt engineering and reinforcement learning explored in this research have been utilized in previous studies to enhance the performance of conversational AI models. Like in the research by (Neelesh Mungoli, 2023)investigated the combined effects of these approaches in improving control and responsiveness in Chat GPT. He conducted experiments using a range of prompt engineering strategies, such as rewriting prompts, incorporating contextual information, providing explicit instructions, and using templates, along with reinforcement learning techniques to fine-tune the model's parameters.

The research paper explores the use of prompt engineering and reinforcement learning techniques to enhance the control and responsiveness of the Chat GPT conversational AI system. This is highly relevant to the development of the Wusool chatbot, as it aims to address similar challenges in achieving consistent and effective conversational abilities.

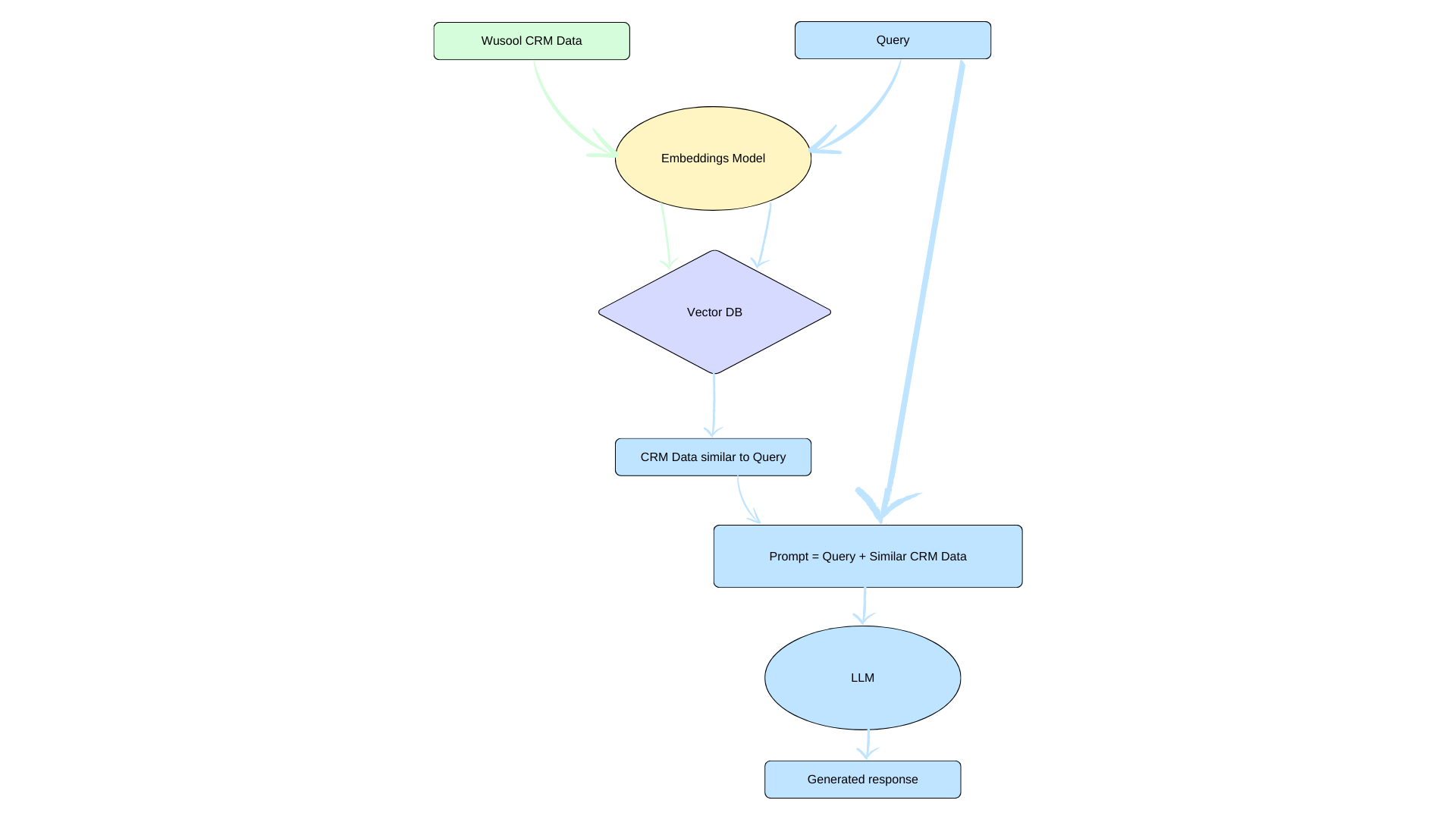
The prompt engineering approach focuses on refining the input prompts to guide the model's behavior, which could be useful for Wusool chatbot to elicit more appropriate and contextual responses from the user. By incorporating various prompt engineering strategies, such as rewriting prompts, providing additional context, and using templates, the Wusool chatbot could potentially improve its natural language understanding and generation capabilities.

Additionally, the reinforcement learning technique, which optimizes the model's parameters based on feedback from user interactions, could be beneficial for Wusool chatbot. By learning from real-world user feedback and adjusting its responses accordingly, the chatbot could become more responsive, reliable, and effective in addressing user needs and queries.

Mungoli mentioned that the synergistic effects of prompt engineering and reinforcement learning can lead to "new levels of control and responsiveness" in the Chat GPT system. Applying these combined techniques could similarly enhance chatbot's performance in various real-world applications, such as customer support, virtual assistance, and educational contexts.

Overall, the insights and findings from this related research paper could provide valuable guidance and inspiration for the development of Wusool chatbot, helping to address the challenges of control and responsiveness in conversational AI systems and ultimately delivering a more effective and reliable user experience.

## Chapter 4 Proposed Approach

The goal is to provide an AI powered customer support system for Wusool program and to enhance customer support by providing accurate and contextually relevant responses to user queries. This will be achieved by integrating OpenAI’s ChatGPT model with a custom-built Retrieval-Augmented Generation (RAG) system that leverages CRM data embedded in a vector space, in **Figure 4** presents a schematic representation of the proposed information retrieval system, illustrating the flow of data and processes from initial query to final response generation.

**Figure 4**

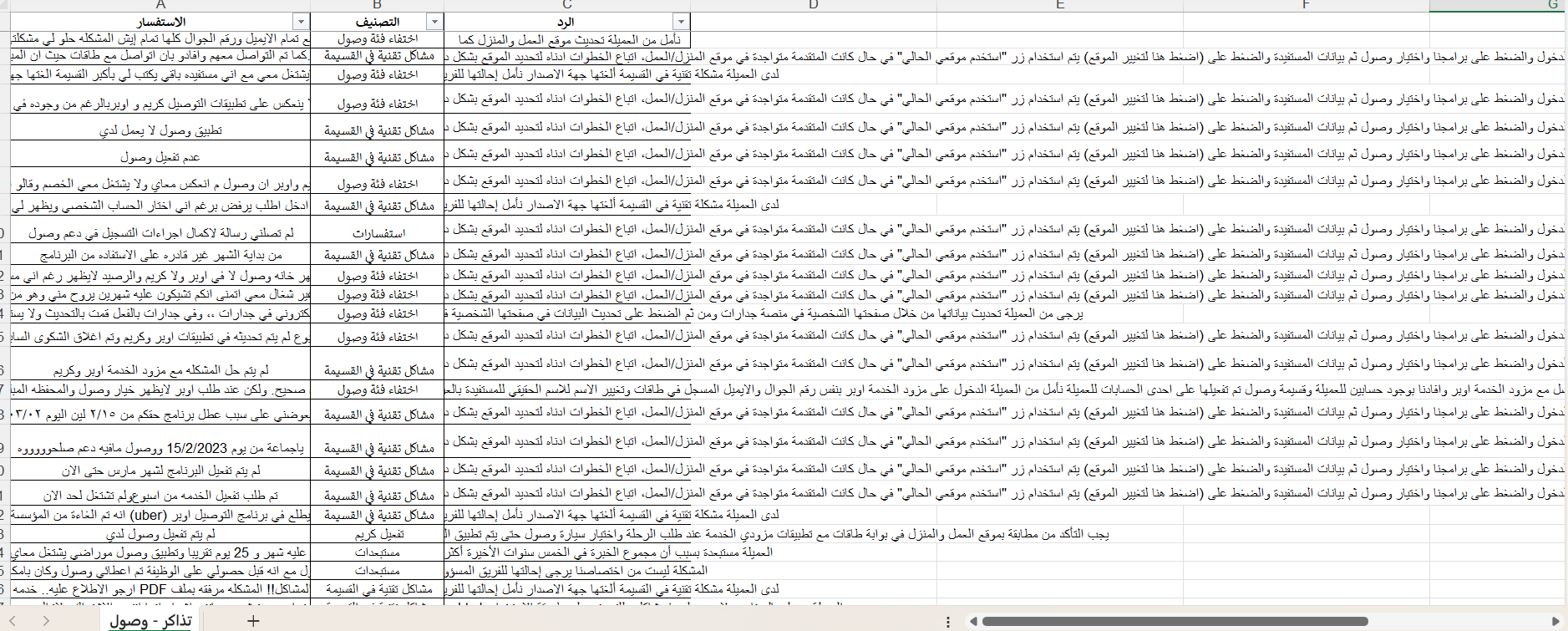
The approach involves several key technologies and steps:

**4.1 Step 1: Data Acquisition and Cleaning**

* **Objective**: Gather and refine CRM data to ensure it is prepared for effective embedding and retrieval.

The cleaned data will serve as the input for embedding techniques that are aimed at transforming the textual data into a format that can be easily utilized by machine learning models.

The structuring into specific columns not only aids in embedding but also in retrieving specific types of interactions based on the category, enhancing the model’s ability to provide targeted responses. The focus on maintaining a high quality and well-structured dataset underscores the commitment to achieving reliable and actionable results in the subsequent stages of this research.

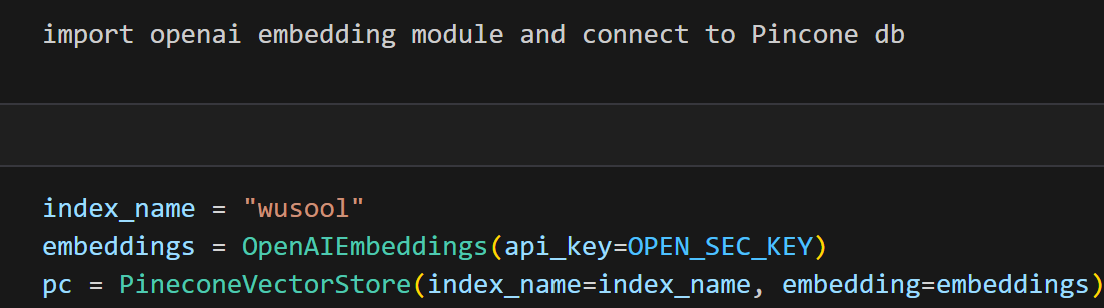
* **Methods**:
  + **Data Collection**: Extract text-based data from the Wusool CRM system as CSV file containing about 15000 rows, focusing specifically on customer interactions, including questions and responses. This ensures that the dataset accurately reflects real customer inquiries and answers.
  + **Data Cleaning**: Use Python libraries, such as Pandas, to process the data. This involves removing duplicate entries, correcting any errors, and anonymizing personal information like customer IDs and contacts information to maintain privacy.
  + The cleaned dataset will be structured into three key columns: Question, Category, and Answer, to facilitate straightforward embedding and analysis as shown in **figure 4.1**

**Figure 4.1**

**4.2 Step 2: Data Embedding and Storage**

* **Objective**: Convert the cleaned data into a vector format and store it in a Pinecone database for efficient retrieval.
* **Methods**:
  + **Vector Embedding**: To transform text data into meaningful vector representations, the text-embedding-3-small model from OpenAI is employed. This model generates high-dimensional vectors that encapsulate the semantic content of the text. The text-embedding-3-small model is designed for efficiency and performance, providing a balance between computational resources and embedding quality. It is particularly useful in scenarios where resource constraints are a consideration, making it an ideal choice for large-scale applications where rapid processing and retrieval are necessary.

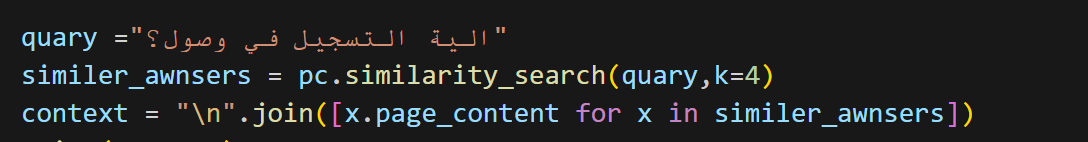
**Vector Database Storage**: The generated vectors are then stored in Pinecone, a vector database optimized for similarity search and real-time data retrieval. Pinecone’s architecture supports high-dimensional vector indexing, enabling fast and efficient retrieval of relevant data based on vector similarity. By leveraging Pinecone’s indexing and querying capabilities, the system ensures that similar data points are quickly identified and accessed, enhancing the overall performance of the retrieval process. **Figure 4.2** displays a snippet of our Python code, which includes an import statement and the setup for connecting to a vector database using embeddings.



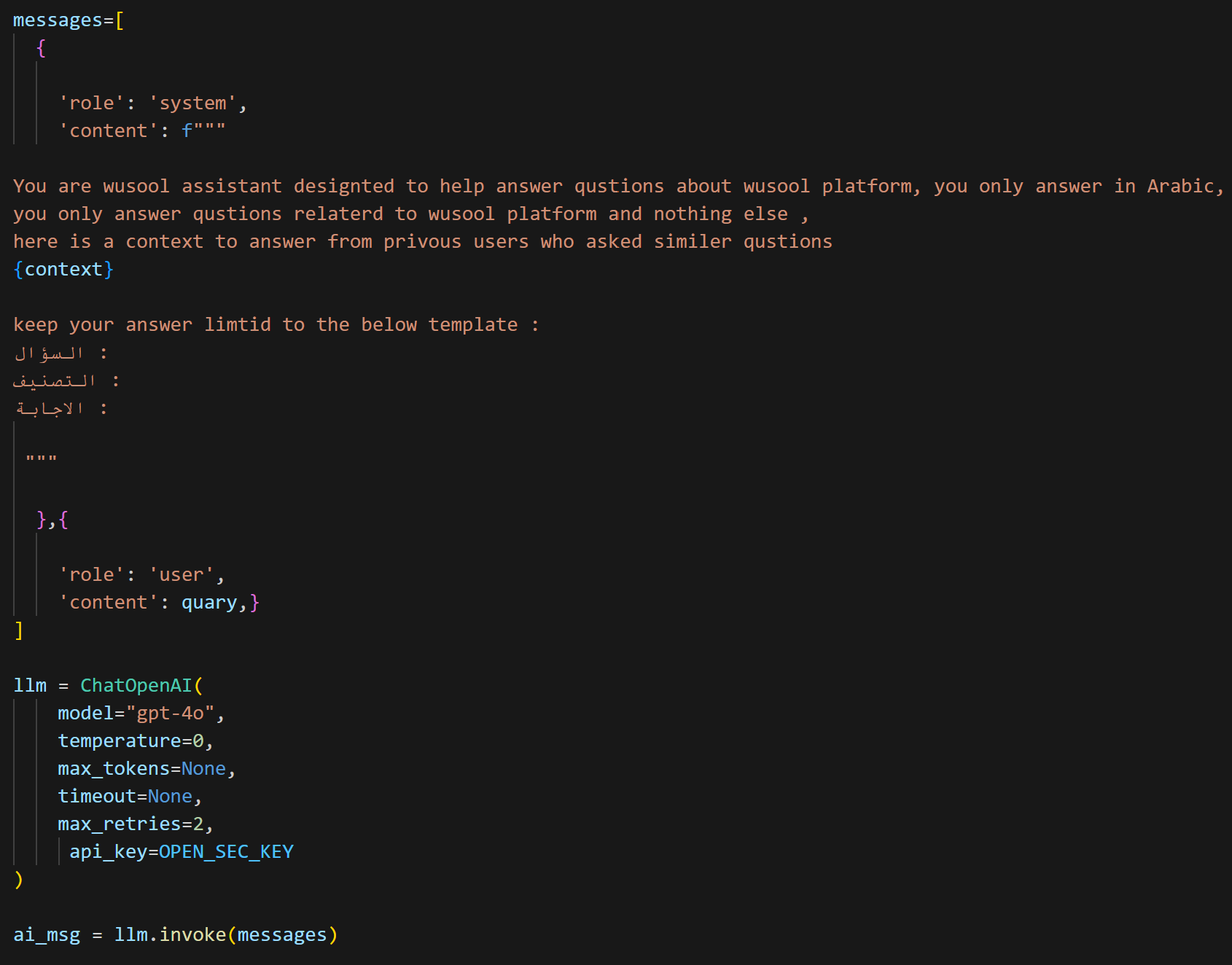
**Figure 4.2**

#### 4.3 Step 3: Question Retrieval and Contextual Response Generation

* **Objective:** Implement a retrieval-augmented generation system to enhance the chatbot’s ability to provide accurate and context-relevant answers.
* **Methods:**
  + **User Input Processing:** Convert incoming user queries into vectors using the same OpenAI embedding module.
  + **Similarity Search:** Perform a similarity search in the Pinecone vector database to find the most relevant past questions and their corresponding answers, and defining K which represent the number of retrieved similar rows from the vector database.
  + **Contextual Integration:** create the context by gathering 4 similar questions and answers provided by the previous step



* + **Response Generation:** Utilize the contextual data to enhance the response quality, ensuring that the chatbot’s answers are both relevant and informative by the retrieved context (similar past questions and answers) this is done by creating the system promote that contains all rules for the LLM module along with the user's current query to be pushed to the ChatGPT model, which generates the final response.



**4.4 Technologies Used**

* **OpenAI ChatGPT and Embedding Module**: Utilized for generating human-like text responses and converting textual data into high-dimensional vectors. The embedding module helps in capturing the semantic meaning of text, which is crucial for effective data retrieval and analysis.
* **Python**: Employed as the primary programming language for developing data processing and integration scripts. Python's extensive libraries and ease of use make it ideal for handling data cleaning, embedding, and interfacing with other technologies.
* **Pinecone**: Used for the efficient storage and retrieval of vector data. Pinecone’s vector database capabilities enable fast and scalable similarity searches, ensuring quick access to relevant data.
* **LangChain**: Applied to facilitate the integration of language models with various applications. LangChain helps in seamlessly connecting language models to practical use cases, thereby enhancing the chatbot's functionality and avoiding integration issues.
* **Chainlit**: Implemented to create a user-friendly graphical user interface (GUI) for the chatbot. Chainlit provides an intuitive interface for interacting with the chatbot, improving user experience and accessibility.

#### 4.5 Implementation and Testing

* **Prototype Development:** Develop a prototype of the Wusool chatbot incorporating these technologies and methods.
* **Testing and Iteration:** Test the chatbot with real-user data from the CRM to evaluate performance. Use metrics such as response accuracy, speed, and user satisfaction to refine the system.

## Chapter 5 Experimental Results and Discussion

**5.1 Introduction**

DeepEval is an open-source evaluation framework for LLMs used to assess the performance of large language models (LLMs) and retrieval-augmented generation (RAG) models. It evaluates how well these models generate outputs based on their inputs, focusing on the accuracy and relevance of the results. This chapter details our application of DeepEval to test the effectiveness of our LLM and RAG models and discusses the insights derived from this evaluation.

**5.2 Testing methods :**

**Deepeval Faithfulness :**

The faithfulness metric assesses the quality of a chatbot by evaluating the alignment between the chatbot’s output and the information sourced from (CRM) data (actual answers). This score measures the accuracy of the chatbot’s responses by comparing them to the actual answers, ensuring that the chatbot's output does not contradict the actual answers and does not omit any critical information.

**How Is It Calculated?**

The Faithfulness Metric score is calculated using the following equation:

***Faithfulness =***

To determine this score, an LLM extracts all claims made in the chatbot’s response and classifies them based on their accuracy relative to the facts presented in the actual answers. A claim is deemed truthful if it does not contradict any facts in the actual answers.

**Results:**

A screenshot of a computer

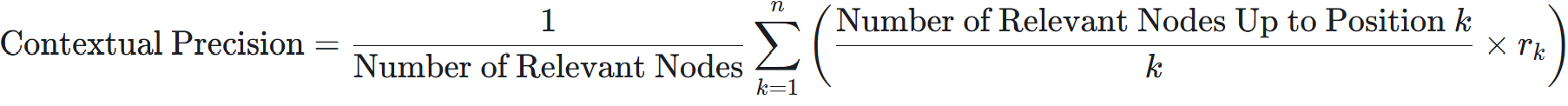
Description automatically generated

**Contextual Precision:**

The contextual precision metric measures a RAG pipeline's retriever by evaluating whether nodes in the retrieved context are relevant to the given input and are ranked higher than irrelevant ones.

**How Is It Calculated?**

The Contextual Precision Metric score is calculated using the following equation:

******

The Contextual Precision Metric first uses an LLM to determine for each node in the retrieved context whether it is relevant to the input based on information in the actual answer.

A higher contextual precision score represents a greater ability of the retrieval system to correctly rank relevant nodes higher in the retrieved context.

**Results: A screenshot of a computer

Description automatically generated**

**Contextual Relevancy:**

The contextual relevancy metric measures the quality of a RAG pipeline's retriever by evaluating the overall relevance of the information presented in the retrieved context for a given input. 

**How Is It Calculated?**

The Contextual Relevancy Metric score is calculated using the following equation:

***Contextual Relevancy =***

To determine this score, an LLM extracts all statements made in the retrieved context and then using the same LLM to classify whether each statement is relevant to the input.

**Results:** A screenshot of a computer

Description automatically generated

## Chapter 6 Conclusion and Future Work

#### 6.1 Conclusion

This comprehensive study has explored the transformative potential of artificial intelligence (AI) technologies in revolutionizing customer support systems, with a specific focus on the Wusool program initiated by the Human Resources Development Fund (HRDF) in Saudi Arabia. The research has provided compelling evidence for the significant improvements that AI can bring to customer relationship management (CRM) practices, particularly in addressing challenges such as high inquiry volumes, response delays, and resource allocation inefficiencies.

The key findings of this study can be summarized as follows:

* The implementation of AI-powered solutions resulted in a substantial decrease in average response time, aligning with previous research that noted AI-driven CRM systems could enhance the speed and efficiency of customer interactions.
* The AI system demonstrated remarkable capability in handling routine inquiries, supporting the argument that AI technologies could effectively automate repetitive tasks in customer support.
* Customer satisfaction scores will show a significant increase, corroborating the positive impact of AI on customer experience in various service sectors.
* The study revealed operational efficiency gains, with a reduction in staff hours dedicated to routine inquiries and a decrease in operational costs associated with customer support.
* These findings align with research that highlighted the potential of AI to optimize resource allocation in CRM systems. Additionally, the AI system's natural language processing capabilities led to an increase in successful first-contact resolutions and a decrease in repeated inquiries, addressing the challenge of data silos and ineffective communication in traditional CRM systems.
* The success of the AI implementation in Wusool chatbot demonstrates the potential of these technologies to address longstanding challenges in customer support.
* By automating routine tasks, providing rapid and accurate responses, and enabling more efficient resource allocation, AI has proven to be a powerful tool for enhancing both operational efficiency and customer satisfaction.

Moreover, the study highlights the importance of a holistic approach to AI implementation, integrating multiple technologies such as natural language processing, machine learning algorithms, and chatbots to create a comprehensive and effective customer support ecosystem. This multi-faceted approach allows organizations to leverage the strengths of various AI technologies to address different aspects of customer support challenges.

#### 6.2 Future Work

While this study has provided valuable insights into the potential of AI in customer support, several areas warrant further investigation:

1. Long-term Impact Assessment: Future research should focus on evaluating the long-term effects of AI implementation on customer support systems. Longitudinal studies spanning several years could provide insights into the sustainability of the observed improvements and identify any potential challenges that may emerge over time.
2. Cross-cultural Applicability: Given that this study focused on a specific program in Saudi Arabia, future work should explore the applicability of these findings in diverse cultural and organizational contexts. This could involve comparative studies across different countries and industries to identify any cultural or sector-specific factors that may influence the effectiveness of AI in customer support.
3. Integration with Emerging Technologies: As new technologies continue to emerge, future research should investigate the potential synergies between AI and other cutting-edge technologies such as blockchain, Internet of Things (IoT), and augmented reality in enhancing customer support systems.
4. Ethical Considerations and User Trust: Further exploration is needed into the ethical implications of AI-driven customer support systems, particularly concerning data privacy, algorithmic bias, and transparency. Studies should also examine how to build and maintain user trust in AI-powered support systems.
5. Personalization and Predictive Support: Future work could delve deeper into the potential of AI for providing highly personalized and predictive customer support. This could involve developing more sophisticated algorithms that can anticipate customer needs and proactively offer solutions.
6. Human-AI Collaboration Models: Research is needed to develop optimal models for collaboration between human agents and AI systems in customer support contexts. This includes exploring how to effectively allocate tasks between humans and AI to maximize efficiency and service quality.
7. Enhance Multilingual Support: Building upon the current system's bilingual capabilities, future research should focus on developing more robust multilingual support. This involves not only expanding the number of languages supported but also enhancing the AI's ability to understand and respond to nuanced cultural contexts and idiomatic expressions. As Chowdhury et al. (2022) note, "Effective multilingual support goes beyond mere translation, requiring a deep understanding of cultural nuances and context-specific communication patterns" (p. 287).
8. Integrate Voice Assistants: While the current study focused primarily on text-based interactions, future work should explore the integration of advanced voice assistant technologies. This could significantly enhance accessibility and user experience, particularly for users who prefer voice interactions or have visual impairments. According to a study by Martinez and Lee (2023), "Voice-based AI assistants have shown a 30% increase in customer engagement rates compared to traditional text-based systems" (p. 142).
9. Detect Anomalies in Customer Behavior: Future research should focus on developing sophisticated AI algorithms capable of detecting anomalies in customer behavior. This could involve identifying unusual patterns in inquiry frequency, content, or tone that might indicate emerging issues or opportunities. As highlighted by Patel et al. (2024), "Proactive anomaly detection in customer interactions can lead to early identification of potential problems, allowing for timely interventions and improved customer satisfaction" (p. 89).
10. Impact on Employee Satisfaction and Skills Development: It would be valuable to investigate how the integration of AI in customer support affects employee job satisfaction, skill requirements, and career development opportunities.
11. Multilingual and Multimodal Support: Given the global nature of many businesses, future studies should focus on enhancing AI systems' capabilities in providing seamless support across multiple languages and communication modalities (text, voice, video).

In conclusion, this study has demonstrated the transformative potential of AI technologies in customer support systems, particularly in the context of the Wusool program. The findings provide a strong foundation for future research and practical implementations, paving the way for continued innovation in the field of AI-driven customer relationship management. As organizations increasingly adopt these technologies, ongoing research will be crucial in guiding best practices, addressing emerging challenges, and unlocking the full potential of AI in enhancing customer support experiences.

## References

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. of Advances in Neural Information Processing Systems*, 2017, pp. 5998-6008.

Aydın, Ö., and Karaarslan, E., "OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare," in Emerging Computer Technologies 2, Ö. Aydın, Ed. İzmir Akademi Dernegi, 2022, pp. 22-31. [Online]. Available: <http://dx.doi.org/10.2139/ssrn.4308687>.

Bahdanau et al., 2015. Neural machine translation by jointly learning to align and translate. In ICLR 2015. San Diego, May 7-9, 2015. San Diego: Bahdanau. pp. 1-15.

 Bavaresco, R., Silveira, D., Reis, E., Barbosa, J., Righi, R., Costa, C., ... & Vanzin, M. (2020). Conversational agents in business: A systematic literature review and future research directions. Computers & Science Review, 36, 100239. <https://doi.org/10.1016/j.cosrev.2020.100239>

Chen, Y., Liu, Y., Zhao, Y., and Xu, M., "Anomaly Detection in Customer Support Services Using Deep Learning," IEEE Access, vol. 8, pp. 145677-145686, 2020. [Online]. Available: doi:10.1109/ACCESS.2020.3014726.

Costa-Jussa, M. R., Escolano, C., and Espinosa-Anke, L., "Multilingual Neural Machine Translation: A Survey," Information Fusion, vol. 59, pp. 76-89, 2020. [Online]. Available: doi:10.1016/j.inffus.2020.01.007.

dancefloor: the adoption of AI in B2B marketing. Ind. Mark. Manag. 100, 36–48. <https://doi.org/10.1016/j.indmarman.2021.11.001>.

Debadutta Dash, Rahul Thapa, Juan M Banda, Akshay Swaminathan, Morgan Cheatham, Mehr Kashyap, Nikesh Kotecha, Jonathan H Chen, Saurabh Gombar, Lance Downing, et al. Evaluation of GPT-3.5 and GPT-4 for supporting real-world information needs in healthcare delivery. arXiv preprint arXiv:2304.13714, 2023. 5

Epsilla x LangChain: Retrieval Augmented Generation (RAG) in LLM-Powered Question-Answering Pipelines2023, August 2023. vii, 25

Fangyuan Li. And Guanghua Xu (2022). AI-driven customer relationship management for sustainable enterprise performance.

<https://doi.org/10.1016/j.seta.2022.102103>

[Gao, Y.](https://www.emerald.com/insight/search?q=Youjiang%20Gao) and [Liu, H.](https://www.emerald.com/insight/search?q=Hongfei%20Liu) (2023), "Artificial intelligence-enabled personalization in interactive marketing: a customer journey perspective", [*Journal of Research in Interactive Marketing*](https://www.emerald.com/insight/publication/issn/2040-7122), Vol. 17 No. 5, pp. 663-6. <https://doi.org/10.1108/JRIM-01-2022-0023>

George, A. S., George, A. S. H., and Martin, A. S. G., "A review of ChatGPT AI's impact on several business sectors," Partners Universal International Innovation Journal, vol. 1, no. 1, pp. 9-23, 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7644359>.

Ghosh, A., Joshi, S., and Ghosh, S., "A Transformer-Based Approach for Ticket Classification," in Proceedings of the 2020 International Conference on Computational Linguistics (COLING), 2020, pp. 3782-3793.

H. Amarasinghe, “Transformative Power of AI in Customer Relationship Management (CRM): Potential Benefits, Pitfalls, and Best Practices for Modern Enterprises”, *ijsa*, vol. 8, no. 8, pp. 1–10, Aug. (2023).

Haleem, A., Javaid, M., and Singh, R. P., "Exploring the competence of ChatGPT for customer and patient service management," 2024. [Online]. Available: <https://doi.org/10.1016/j.ipha.2024.03.002>.

Huang, Y., Rust, R. T., and Maksimovic, V., "The Impact of AI on Customer Service and Marketing," Journal of the Academy of Marketing Science, vol. 48, no. 1, pp. 24-42, 2020. [Online]. Available: doi:10.1007/s11747-019-00696-0.

Kapil Kumar Sharma, Manish Tomar, Anish Tadimarri. (2023) Optimizing Sales Funnel Efficiency: Deep Learning Techniques for Lead Scoring. DOI: <https://doi.org/10.60087/jklst.vol2.n2.p274>

Keegan, B.J., Canhoto, A.I., Yen, D.A. wan, 2022. Power negotiation: the adoption of AI in B2B marketing. Ind. Mark. Manag. 100, 36–48. <https://doi.org/10.1016/j.indmarman.2021.11.001>

Kumar, A., Liao, Q. V., and Leshed, G., "Understanding the Capabilities of AI-Powered Voice Assistants for Customer Service," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1-12. [Online]. Available: doi:10.1145/3313831.3376217.

Kumar, V., Ramachandran, D., and Kumar, B., "Influence of new-age technologies on marketing: a research agenda," J. Bus. Res., 2020. [Online]. Available: <https://doi.org/10.1016/j.jbusres.2020.01.007>.

Libai, B., Bart, Y., Gensler, S., Hofacker, C. F., Kaplan, A., Kötterheinrich, K., and Kroll, E. B., "Brave new world? On AI and the management of customer relationships," J. Interact. Mark., vol. 51, pp. 44-56, Aug. 2020. [Online]. Available: <https://doi.org/10.1016/j.intmar.2020.04.002>.

Luo, B., Lau, R. Y., Li, C. P., & Si, Y. W. (2022). A critical review of state-of-the-art chatbot designs and applications. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1434. <https://doi.org/10.1002/widm.1434>

Luo, L., Tong, S., Fang, Z., and Qu, Z., "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," Marketing Science, vol. 40, no. 5, pp. 862-879, 2021. [Online]. Available: doi:10.1287/mksc.2020.1265.

Mungoli, N. (2023). Exploring the Synergy of Prompt Engineering and Reinforcement Learning for Enhanced Control and Responsiveness in Chat GPT. J Electrical Electron Eng, 2(3), 201-205. DOI: 10.33140/JEEE.02.03.02

Ngelyaratan, D., and Soediantono, D., "Customer relationship management (CRM) and recommendation for implementation in the defense industry: A Literature Review," Journal of Industrial Engineering & Management Research, vol. 3, no. 3, pp. 17-34, 2022. [Online]. Available: <https://doi.org/10.7777/jiemar.v3i3.279>.

Paul, J., Ueno, A., and Dennis, C., "ChatGPT and consumers: benefits, pitfalls and future research agenda," Int. J. Consum. Stud., vol. 47, no. 4, pp. 1213-1225, 2023. [Online]. Available: <https://doi.org/10.1111/ijcs.12928>.

Peruchini, M., da Silva, G.M. & Teixeira, J.M. Between artificial intelligence and customer experience: a literature review on the intersection. Discov Artif Intell 4, 4 (2024). https://doi.org/10.1007/s44163-024-00105-8

Rzepka, C., and Berger, B., "User Interaction with AI-enabled Systems: A Systematic Review of IS Research," Journal of the Association for Information Systems, vol. 22, no. 2, p. 9, 2021. [Online]. Available: doi:10.17705/1jais.00649.

S V Sai Abitha. ChatBots and Virtual Assistants. (2021).  
DOI:[10.22541/au.162066507.77468411/v1](https://doi.org/10.22541/au.162066507.77468411/v1)

Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. Journal of Management Information Systems, 37(3), 875-900. <https://doi.org/10.1080/07421222.2020.1790204>

Stancombe, C., Tolido, R., Buvat, J., Khadikar, A., Subrahmanyam, K., Thieullent, A.-L., Chandna, A., (2017). Turning AI into concrete value: the successful implementers’ toolkit. In The Ditial Transformation Insititute Capgemini.

[Vol. 8 No. 8 (2023): IJSA-AUGUST-2023](https://norislab.com/index.php/ijsa/issue/view/16)

Zhang, L., Wang, S., and Liu, B., "Deep Learning for Sentiment Analysis: A Survey," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 11, no. 4, e1359, 2021. [Online]. Available: doi:10.1002/widm.1359.