A Dynamic Query Framework for Research Accessibility using OpenAI and Langchain

Louise Gabrielle L. Talip and Reginald Neil C. Recario

I. Introduction

In the ever-evolving realm of education and academic research, the need for innovative tools and technologies remains a persistent challenge. The rise of the digital age gave way to a wide variety of tools such as for data collection and analysis (e.g. Google Sheets, Python), collaborative writing (e.g. Overleaf), and even for plagiarism detection (e.g. Turnitin). However, these tools primarily cater to experienced researchers because they often require a level of technical expertise that can be challenging for individuals without specialized knowledge. Furthermore, while these tools excel in writing and coding tasks, they fall short in facilitating the crucial aspect of comprehending existing research literature — a cornerstone of effective research. Reading is a crucial aspect of the research process, influencing both topic selection and formulation [1]. It is a fundamental step in shaping one's own research direction and synthesizing new knowledge as it is not just a supplement to empirical methods, but a method of inquiry in itself [2]. By comprehending previous studies, researchers can actively shape the path of their own investigations. Currently, there remains a scarcity of resources for these researchers, particularly students, in initiating their academic journey, especially when it requires specialized skills and deeper technical and theoretical understanding. This gap in accessible tools and resources becomes especially apparent in countries like the Philippines, where the pursuit of research is gradually declining [3] [4]. It is imperative to foster the culture of innovation and research among scholars through accessibility and approachability. Developing entry-level tools like a research assistant chatbot tailored for academics in the Philippines could serve to bridge this gap. By allowing access to simplified academic resources, there is potential to stimulate increased interest and participation in research, potentially contributing to the nation's intellectual advancement and innovation.

A. Background of the Study

Chatbots are intelligent systems that can hold conversations with humans using natural language [5]. They have a wide range of applications, including education, information retrieval, business, and e-commerce [6]. Over the years, chatbots have been developed to automate interactions, streamlining communication between users and providing convenient access to valuable information for both businesses and individuals. They prove particularly useful in industries where immediate assistance might not be available, serving as an accessible resource. Chatbots efficiently address common inquiries such as specific restaurant operating hours, detailed airline guidelines, or government institution procedures.

The use of chatbots in education has been on the rise in recent years [7]. These AI-powered virtual assistants have demonstrated significant potential in streamlining various tasks such as student onboarding or providing instant general support to students such as Dashly and enhancing the learning experience such as Socratic, Mathway, and Duolingo. Chatbots have become integral in improving student engagement, retention, and overall educational outcomes, especially now with the rise of ChatGPT and applications powered by it.

Chatbots either execute specific tasks or engage in open-ended conversations. Many methodologies have emerged in this domain, spanning from initial hard-coded response systems to more sophisticated Artificial Intelligence-driven advancements, showcasing the evolution of techniques employed in their development [8]. With the advancements in artificial intelligence, chatbots have progressed to offer more comprehensive and dynamic answers. They can now be trained or prompted to understand a wider scope of topics with the aid of appropriate datasets, such as in the case of OpenAI's ChatGPT [9]. Consequently, diverse sectors, including education, have embraced these evolving technologies.

The rise of Large Language Models have been revolutionizing the landscape of natural language processing. Their ability to comprehend and generate text with a human-like fluency has led to their wide adoption across various industries [10], including education. LLMs have become indispensable tools, aiding students, educators, and researchers in a variety of tasks. They can now assist in drafting essays and automating data analysis such as ChatGPT [9] or writing code through GitHub's Copilot [11].

Due to the popularity and demand of these models, developers have also created frameworks to better automate processes and build applications on top of the OpenAI API. One of these tools is called Langchain, a framework specifically designed to facilitate the creation of said applications. Its objective is to enable developers to use any large language model and use it across multiple sources of data or seamlessly interact with multiple APIs [12].

Presented to the Faculty of the Institute of Computer Science, University of the Philippines Los Baños in partial fulfillment of the requirements for the Degree of Bachelor of Science in Computer Science

This paper explores the potential and significance of developing a chatbot-driven solution aimed at addressing inquiries and providing recommendations related to academic papers. Its primary focus lies in aspiring researchers in the Institute of Computer Science at the University of the Philippines Los Baños. This system utilizes OpenAI's advanced language model functionalities integrated with the Langchain framework, coupled with Chainlit for its user interface.

B. Statement of the Problem

Guido and Mangali [13] state that "a country's economic and intellectual wealth is associated with the productivity of its research". Their study indicated a notable connection between the scores from the Programme for International Student Assessment (PISA) and a country's research productivity. The Philippines in particular has been found to have a low research output and low PISA scores. The Philippines falls below the average band line for Southeast Asian nations in terms of published documents, citations, self-citations, and open access [14]. Despite the country recording the lowest PISA scores among its Southeast Asian counterparts, its research performance and quality remained relatively commendable within the region. In a study in 2015 by Wa-Mbaleka [15], one of the major factors that lead to limited publications is the time constraints. Reviewing hundreds of pages of papers is a notably time-intensive task. This is true for both new and seasoned researchers – where the former would be overwhelmed and the latter, which mostly consist of professors with hectic schedules, do not have the luxury to dedicate a lot of their time to sift through many pages. A recent study by Lobo [16] unraveled that students have encountered obstacles related to infrastructure, communication, and time management while undertaking their thesis work. Bueno [4] in 2019 also confirmed Wa-Mbaleka's findings in their study, but also added that the lack of solid foundation, along with the lack of resources, also contributes to the lack of interest from researchers to publish more papers. With these factors and many more, it has been increasingly difficult to find funding, which has always been a major barrier to research. The 2015 report from the UNESCO Institute of Statistics [17] [18] highlights that a mere 0.2% of the country's gross domestic expenditure is directed toward research and development, which is significantly lower than the global average. This negative cycle continues today where the minimal return in the country's research output, the lack of research culture, and the lack of time leads to a withdrawal or reallocation of funding from the government, which in turn either discourages researchers or pushes them to pursue their career in other countries.

The Philippines holds considerable promise for the development of research in various fields, creating significant strides despite the limited resources such as the development of the country's own COVID-19 test kits [19]. Moreover, there are over 7,000 published researchers from different backgrounds and fields in the country according to the ADS Scientific Index [20] and there is also a steady increase in the number of publications since the early 2000s, according to Scival [21] [22]. Despite the other factors that hinder research development, it is still important to contribute and aid in the plight of advancing Philippine research by making it a little bit easier for researchers in one way or another.

C. Objectives of the Study

The primary focus of this study is to create a research assistant chatbot to aid researchers in understanding scientific publications or other scholarly articles in the Institute of Computer Science of the University of the Philippines Los Baños.

Specifically, this study aims to:

1) Analyze and implement Langchain and OpenAI to enhance the chatbot's conversational understanding capabilities in complex academic contexts, addressing the challenge of natural language processing.
2) Utilize and employ Langchain, OpenAI, and Chainlit functionalities to devise efficient methods for content extraction from thesis papers, focusing on the retrieval of correct and simplified answers, creating accessible summaries, keywords, and citations to help researchers comprehend academic content effectively.
3) Evaluate and improve the accuracy of information retrieval from academic papers with Chainlit's Feedback feature.
4) Assess the chatbot's effectiveness through Post-Task Evaluation survey and System Usability Scale (SUS).

D. Significance of the Study

The research and academic landscape in the Philippines are relatively niche, deterring many potential students from pursuing these paths due to limited experience and a steep learning curve. Addressing this, the development of a tool becomes pivotal in revitalizing researchers' interest.

Enhancing the research flow led to advanced chatbots that surpass mere keyword matching, now capable of understanding the context behind user queries. Existing applications like ChatPDF [23] and the AI Research Assistant by Elicit [24] exhibit limitations, handling only one PDF at a time and having restricted datasets, often excluding Philippines-based papers. In contrast, a tool incorporating custom data offers accessibility to local resources in the university library.

High costs associated with proprietary models hinder accessibility, especially for entry-level scholars with limited purchasing power. Utilizing Langchain technology improves scalability by dividing tasks into smaller subtasks, saving tokens for data retrieval and expediting the process. This approach, akin to data caching, optimizes efficiency by targeting specific segments containing required information. In the Philippines, where GPU demand is high, tools like Langchain become crucial for model training, enhancing accessibility even when using existing models like OpenAI. Notably, the cost of existing applications disproportionately affects entry-level scholars.

Simplifying the research workflow accelerates the identification of relevant papers and information, expediting the writing process. Improved accessibility to knowledge fosters

a deeper understanding of subjects, particularly for emerging researchers who can ask follow-up questions, thereby lowering barriers and inspiring greater participation in the research process.

II. REVIEW OF RELATED LITERATURE

A. Evolution of Chatbot Technology

The evolution of chatbot technology has been significant, with a shift from rudimentary models to advanced intelligent systems [25]. The term "Turing Test" emerged from Alan Turing's research paper entitled "Computing Machinery and Intelligence" in the 1950s where an artificial intelligence has to pass as human according to certain benchmarks. The development of ELIZA in the Massachusetts Institute of Technology in 1996 led to the first chatbot that passed the Turing Test due to how they didn't foresee how many individuals would easily attribute human-like emotions to the program. According to Adamopouloou [26], it employed basic pattern matching and a response system based on templates. Deshpande further states that the first few chatbots operated on these predefined response pairs linked to specific inputs. Using pattern matching and string processing, it facilitated ongoing conversations between computers and humans, but it lacked the criteria to be considered an intelligent chatbot. Chatbots have then evolved from simple keyword recognition to "evaluating user input victimization through heuristical pattern matching rules" such as in the case of ALICE (Artificial Linguistic internet computer Entity) in 1995. This evolution has been driven by the maturation of AI technologies, the integration of NLP, and the recent developments of using Recurrent Neural Network (RNN) and Long Short Term Memory (LSTM) to achieve AI [27]. The prevalence of chatbots has been on the rise, finding applications across diverse fields such as marketing, education, and healthcare [26], which is clear in Figure 1 below – Scopus search results from 2000 to 2019 using keywords like "chatbot," "conversation agent," or "conversational interface".

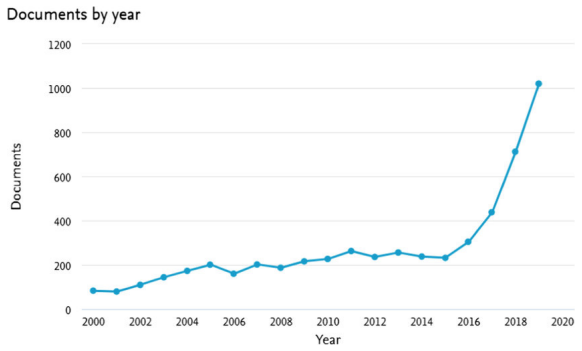


Fig. 1: Scopus search results (2000-2019) of the keywords "chatbot" or "conversation agent" or "conversational interface" [26]

Presently, the release of OpenAI's ChatGPT has sparked significant interest and concern in the AI community. Haque in 2023 [28] emphasizes the revolutionary nature of ChatGPT and its potential to automate a range of tasks such as translation, customer service, and content generation could significantly influence and dominate industries because of its ability to generate human-like responses and answer a broader range of topics. This is because it has been trained with a model that has 175 billion parameters allowing it to generate quality responses.

B. Language Models and OpenAI

Large Language Models (LLMs) are a groundbreaking development in the field of artificial intelligence as they are a class of language models that have demonstrated exceptional performance in natural language processing tasks both in understanding and generation [29]. Moreover, LLMs are pre-trained on a vast database of text and possess the ability to generate coherent and contextually relevant text in response to user queries. These models, armed with their extensive knowledge, serve as the foundational technology for many chatbot systems.

According to MindsDB [30], OpenAI, a pioneering research organization in artificial intelligence, has played a pivotal role in advancing LLMs and making them accessible to the broader community. Their GPT-3.5 model, in particular, has garnered widespread attention for its exceptional language generation capabilities. OpenAI's infrastructure and API have enabled developers to create sophisticated chatbots and applications that leverage the power of GPT-3.5. Now that they have released GPT-4, it has been the leading model in terms of quality responses. MindsDB [30] also shows in the table below the models that are at the forefront today.

Model	Provider	Open-Source	Speed	Quality	Params	Fine-Tuneability
gpt-4	OpenAI	No	★☆☆	★★★★	-	No
gpt-3.5-turbo	OpenAI	No	★★☆	★★★★	175B	No
gpt-3	OpenAI	No	★☆☆	★★★★	175B	No
ada, babbage, curie	OpenAI	No	★★★	★☆☆☆	350M - 7B	Yes
claude	Anthropic	No	★★☆	★★★★	52B	No
claude-instant	Anthropic	No	★★★	★★☆☆	52B	No
command-xlarge	Cohere	No	★★☆	★★☆☆	50B	Yes
command-medium	Cohere	No	★★★	★☆☆☆	6B	Yes
BERT	Google	Yes	★★★	★☆☆☆	345M	Yes
T5	Google	Yes	★★☆	★☆☆☆	11B	Yes
PaLM	Google	Yes	★☆☆	★★☆☆	540B	Yes
LLaMA	Meta AI	Yes	★★☆	★★☆☆	65B	Yes
CTRL	Salesforce	Yes	★★★	★☆☆☆	1.6B	Yes
Dolly 2.0	Databricks	Yes	★★☆	★★☆☆	12B	Yes

Fig. 2: Table of Large Language Models [30]

Because of this, OpenAI is the optimal model to use for the chatbot to ensure the best responses for each query. Depending on the use case of the prompt, a developer can choose from the different models that OpenAI offers – some models are more appropriate for conversational chats while others are suitable for direct answers to queries. The models are also priced differently due to the number of tokens that it can accept and generate. The tables below from the official OpenAI website shows the differences of these models [9].

Model Name	Tokens	Training Dataset
gpt-4-1106-preview	128,000	Up to Apr 2023
gpt-3.5-turbo-1106	16,385	Up to Sep 2021
gpt-3.5-turbo	4,096	Up to Sep 2021
gpt-3.5-turbo-16k	16,385	Up to Sep 2021
gpt-3.5-turbo-instruct	4,096	Up to Sep 2021

TABLE I: Comparison of OpenAI Models - Tokens and Training Dataset [9]

Model Name	Pricing (Input)	Pricing (Output)
gpt-4-1106-preview	$0.01 / 1K tokens	$0.03 / 1K tokens
gpt-3.5-turbo-1106	$0.0010 / 1K tokens	$0.0020 / 1K tokens
gpt-3.5-turbo	$0.002/1K tokens	$0.002/1K tokens
gpt-3.5-turbo-16k	$0.003 /1K tokens	$0.004 /1K tokens
gpt-3.5-turbo-instruct	$0.0015 / 1K tokens	$0.0020 / 1K tokens

TABLE II: Comparison of OpenAI Models - Pricing for Input and Output [9]

C. Applications in Academia and Research

According to Shawar and Atwell [6], originally, chatbots were created for amusement and to simulate human dialogue. While this purpose remains prevalent in chatbot development, the increased traction of this technology has led to diverse applications. Subsequently, chatbot technology has found utility in various domains, including information retrieval, query resolution, facilitating evidence-based decisions, and many more. Abdelhamid [31] further explores the use of chatbots as smart teaching assistants, stating that students are much more likely to use and engage with the system than ask questions directly to their professors.

The prevalence of chatbots in education and research has also been increasing, especially now with the advancement of language models [32]. Conversational agents, which facilitate easier access to information through engaging interactions, have prompted tools to improve development processes. This includes refining user experience and design, improving frameworks and platforms, among other enhancements.

However, the integration of chatbots in education has been focused on skills such as subject-learning, reading and writing, speaking, and language translation [33]. The utilization of OpenAI's ChatGPT has significantly influenced how learners process and comprehend information. The results in Memarian's [34] study highlighted that ChatGPT offers extensive potential for various applications, including personalized and complex learning, customized teaching approaches, diverse learning activities, assessments, asynchronous communication, feedback systems, accurate research, personas, task distribution, and cognitive support.

There are still gaps and a lack of tools, especially in the field of research. Kooli [7] stated that AI technologies are envisioned to revolutionize research and education by automating laborious and repetitive tasks, assisting in data analysis, and fostering innovative modes of learning and evaluation. AI-powered research assistants would be advantageous in the academic field as it helps in fact-checking and offers quick access to pertinent information. These systems serve as efficient tools for gathering data, operating round-the-clock and processing large volumes of data swiftly. This capability enables researchers to amass high-quality data and obtain precise information at any time, thereby mitigating the potential for human error. Nonetheless, Kooli emphasizes that the primary accountability for the content and quality of the research still remains with the human researcher. A research chatbot functions solely as a tool, not as a substitute for the indispensable role of the researcher. Similar to human research assistants, these chatbots need thorough and ongoing oversight to prevent deviations.

Lund and Wang [35] states that ChatGPT holds potential for improving the academe in multiple ways, such as its advanced capabilities in literature reviews, generating text, automating summarization, and answering queries. It can also assist through extensive search and referencing features. These capabilities offer researchers the opportunity to streamline their workload, allocating more time and energy towards the creative and analytical facets of their endeavors. However, constraints within ChatGPT and the absence of an effective framework pose limitations. Hence, the creation of a structured chatbot using a framework like Langchain stands as an ideal approach to manifest these capabilities.

Lastly, according to Klimova and Seraj [33], AI-powered chatbots significantly influence students' skills with the use of text, speech, graphics, haptics, and gestures to assist them in completing their tasks. These chatbots extend a broader spectrum of services to students, enhancing their motivation and expanding the traditional learning paradigm into the digital realm. This offers a promising outlook, suggesting the potential to augment research processes, facilitating an easier, more accessible way of learning and fostering an increased interest in the field.

D. Langchain

Topsakal and Akinci [12] defines LangChain as a framework, developed by Harrison Chase, that is designed for creating applications that leverage a variety of large language models. It serves as a middleware layer, where its primary objective is to facilitate developers in seamlessly integrating various data sources and APIs from a diverse range of applications. To achieve this, LangChain offers components, which are modular abstractions and chains – adaptable pipelines customized for specific use cases. Sreeram and Sai in 2023 [36] adds that specifically, it is "a cutting-edge solution which helps us in the querying process and extracting information from PDFs. With its advanced NLP algorithms, it helps users to interact with the PDFs and makes the document search and retrieval very easy". Langchain's unique capabilities make it a valuable asset in creating a chatbot that can provide users with accurate and efficient access to different sources of data, including research papers.

The primary components from Langchain that will be used are "chains" and "agents". As stated in the official Langchain documentation [37], chains are the building blocks of Langchain, allowing the developers to generate responses from the chosen LLM using templates and the user input. Topsakal and Akinci [12] explains that these chains can be structured and sequenced in ways beyond the capabilities of standard LLM APIs. For instance, they can utilize preceding chain responses as inputs for subsequent chains. These

chains may be organized linearly or amalgamate answers from multiple chains into a single conclusive chain. Additionally, some chains can be routed to different chains with varying templates based on user input. Agents, on the other hand, are the decision makers. Agents hold a collection of tools curated for specific usage based on the user input. These tools can contain functions utilizing Python code, more chains, or different APIs and libraries, enabling an extensive scope of capabilities. Additionally, agents have the capacity to call other agents, initiating additional sequences of actions.

E. Limitations and Challenges

The use of tools can only do so much for Philippine researchers, who face a multitude of challenges. These include limited funding and resources [38], exacerbated by issues of accessibility and local attitudes towards science [39]. Furthermore, academic staff in the country are often ill-prepared for their roles in research training and supervision [40] and the educational system as a whole is plagued by perennial issues [41]. These challenges collectively hinder the effectiveness of tools in supporting the research efforts of Philippine scholars. While a chatbot system would aid in simplifying the research process, there is no guarantee of the gravity of its impacts to the research community as a whole and its influence on the culture and perspective towards research in general.

Although the implementation of an AI chatbot would be advantageous and convenient within an ideal context, it is essential to acknowledge that there are also inherent limitations concerning its functionalities and the implications for its potential users. Some AI models tend to be biased in a number of complex ways, mostly stemming from various factors including training data, model specifications, and algorithmic constraints, product design, and policy decision [42]. A study by Cirillo, et. al. in 2020 [43] has proved that the majority of present biomedical AI technologies lack mechanisms to detect biases and that its design fails to consider the differences in health and disease outcomes influenced by the sex and gender of an individual, which will heavily alter the way it will suggest treatment and holds the potential to "produce mistakes or discriminatory outcomes". ChatGPT, for example, has a tendency to generate biased responses. The various shortcomings encompass superficial, inaccurate, or incorrect content, as stated in a study by Sallam in 2023 [44]. Ethical concerns and risks emerged that are associated with biased data in training and potential plagiarism concerns. A crucial aspect mentioned in the study as well was the concept of ChatGPT hallucination, a risky phenomenon that demands careful evaluation by experts or researchers. This is particularly significant due to ChatGPT's capacity to produce scientifically plausible yet incorrect content, emphasizing the necessity for vigilant scrutiny and evaluation within these domains. There have also been a number of case studies that highlight the citation inaccuracies from ChatGPT, as well as insufficient or non-existing references [45]. These are concerning, as biases can influence the information and opinions it generates, potentially perpetuating harmful language patterns [46].

Another drawback introduced by Lopez and Qamber in 2022 [47] is a lack of human assistance in situations where the questions are too complex for the chatbot to answer or when the user prefers human contact. Salvagno [48] claims that AI chatbots such as ChatGPT possess the potential to support the writing process of scientific papers and aid in conducting literature reviews, pinpointing research inquiries, presenting an overview of a field's current status, and providing assistance in tasks like formatting and language review. Both studies confirm that chatbots can serve as tools that aid human researchers; however, they should not substitute for the expertise, judgment, and individuality of human researchers.

LLMs have shown promise in natural language processing and its integration into educational chatbots remains a multi-faceted challenge. It is evident that there is a need for tools that can help aspiring researchers effectively extract relevant content from papers to alleviate the challenges associated with writing and to foster an environment that sparks their interest in research writing.

III. METHODOLOGY

A. Theoretical Framework

The theoretical framework for this chatbot primarily relies on the capabilities of Langchain. It provides the components for the data ingestion and data processing. User interaction is provided by the Python library Chainlit and the large language model to be used is OpenAI's GPT-4 Turbo. The framework starts with the preprocessing phase, where the chatbot ingests and organizes the dataset provided. This preparatory stage involves data collection, PDF loading, text chunking and splitting, and storage in a structured format, establishing a robust foundation for subsequent user interactions. The operational phase focuses more on user queries, intelligent decision-making agents, and information retrieval mechanisms. This framework merges contemporary techniques such as that of Topsakal and Akinci in 2023 [12] and Sreeram [36] for building applications using large language models, enabling the chatbot to adeptly handle a wide array of user queries.

1) Preprocessing Phase: Prior to the chatbot's operational phase, a crucial preprocessing stage takes place in order to prepare the data to be used by the chatbot to answer user queries. The data ingestion involves several key tasks:

1) **Data Collection:** A systematic gathering of relevant data sources, specifically the Special Problems and Thesis manuscripts from the Institute of Computer Science. This collection will focus on documents in PDF format retrieved from the ICS Library archives spanning from January 2018 to December 2023, ensuring adherence to Data Privacy limitations and institutional regulations.
2) **PDF Chunking and Text Splitting:** The collected PDFs will be loaded using Langchain in order to undergo a process of segmentation into manageable sections or 'chunks'. This chunking process involves breaking down lengthy documents into smaller, more digestible sections, enabling efficient handling and analysis. Post-segmentation, these chunks will undergo parsing to extract essential textual content, subsequently stored as 'embeddings'. These embeddings serve as numerical

representations of the text, facilitating streamlined retrieval and comparison. Each chunk is assigned numerical data based on its content relevance, aiding later retrieval of closely related chunks upon query.

3) **Database Storing:** Once the embeddings are generated, the extracted text segments are converted into structured formats, where each embedding has its respective index, and is stored within a vector store or database such as Pinecone. Storing these text segments in a vector store is crucial to optimize the efficiency of the retrieval, particularly during the operational phase of the chatbot. By storing the textual information in this manner, the chatbot can swiftly access relevant content, enhancing its responsiveness in providing accurate and timely responses to user queries.

2) Operational Phase - Chatbot Functionality: The operational phase involves user interaction with the chatbot, in which the process is as follows:

1) **User Selection of Topic and Query:** The chatbot prompts the user to select a specific topic of interest. Then, the user poses their query or question related to the chosen topic.
2) **Langchain Agents and Tools:** A Langchain agent, equipped with decision-making capabilities, determines the appropriate tool to employ and dataset to access for acquiring the proper response. There is no limit to the number of tools that an agent can have, but one needs to be careful with the prompts as sometimes having too many tools with convoluted instructions could also make the agent hallucinate. This decision-making process entails accessing the correct embeddings or representations within the database corresponding to the user's query.
3) **Information Retrieval:** Leveraging the ingested data from the preprocessing phase, the chatbot retrieves relevant data shards from the previously stored documents. These data shards are ranked by the OpenAI embeddings depending on how related they are to what the user is querying for, enabling the extraction of accurate and contextually appropriate information from diverse PDF sources.
4) **Formulation of Response:** The information retrieved from the embeddings is then passed to the LLM. It processes this data and formulates a comprehensive, properly formatted response tailored to address the user's query. This response is subsequently communicated back to the user, completing the interaction cycle.

Refer to Appendix I for a detailed explanation and visual representation of the framework's functionality.

B. System Flow

The system will simulate the process of a text-based conversation between the user and the chatbot application. The user will be able to send queries to the web application about a topic within the scope of the dataset – in this case, the research papers from the Institute of Computer Science from the University of the Philippines Los Baños.

In this flow, the chatbot is designed to provide tailored responses based on the nature of the user's query. This system differentiates between two types of inquiries:

1) **Direct Answer Queries:** These questions seek specific, concise information that is retrieved from a research paper or a collection of papers. In response, the chatbot engages in a conversational manner, delivering the requested information in a user-friendly format. Refer to Appendix II-A for a visual representation of this feature.
2) **List-based Queries:** These encompass inquiries requesting lists of papers related to the query, such as general topics, different technological stacks, or authors of research papers. In these cases, the chatbot responds with a list of research paper titles and the relevant information related to the queried topic. Refer to Appendix II-B for a visual representation of this feature.

C. Website Development

The chatbot will be accessible through a dedicated web application designed to facilitate user interaction. This study will utilize Chainlit, a robust Python library known for its efficiency in constructing user interfaces, to provide the design and to create the essential website components, including a straightforward textbox for user input, separate text boxes for user and AI responses, and interactive buttons to enhance user interaction. The website's core focus lies in the underlying technology empowering the chatbot. Therefore, this study only requires and will predominantly focus on this singular page.

D. Agents and Tools

The use of Agents is the core of this application. According to the Langchain documentation [37], the idea is that the agent will use a language model to choose a sequence of actions to take. It will contain tools and accept a prompt. The tools are simply functions that an agent can invoke if it decides that the specific tool is required to answer a query. In creating a tool, it needs to be described in a way that is most helpful to the agent. This will require rigorous testing so that the agent knows what to do with a particular tool. Refer to Appendix III for the visual representation of the agent framework.

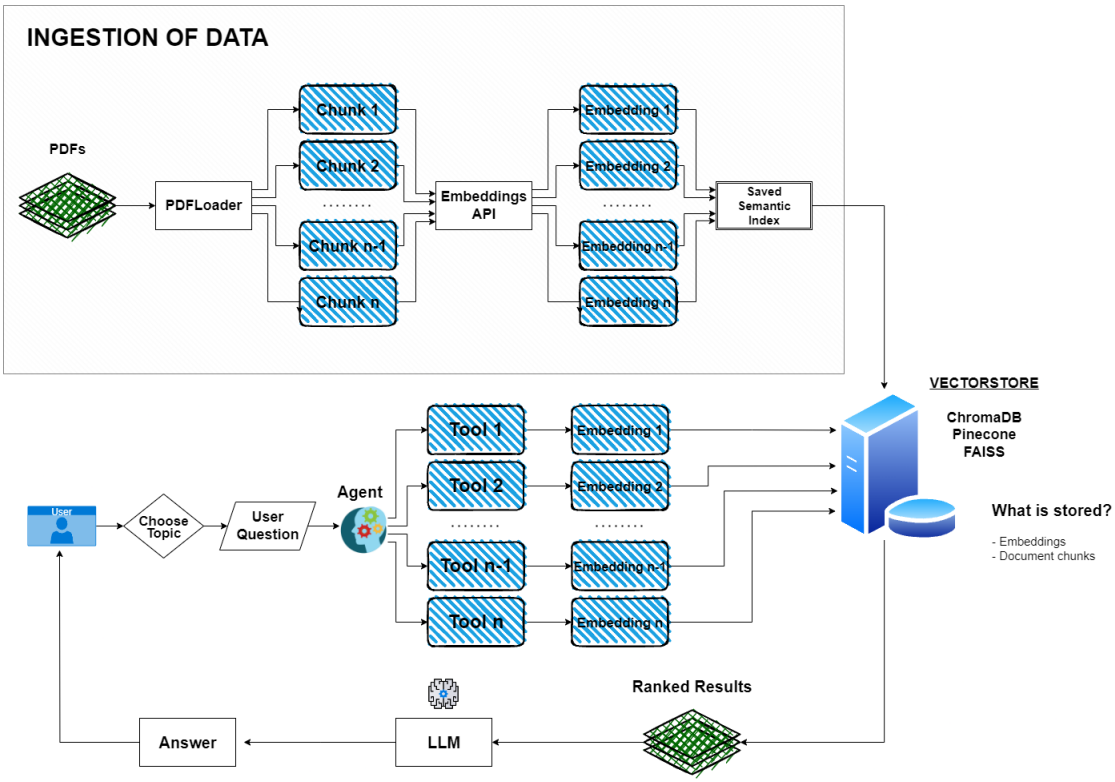
E. Prompt Engineering

Gu [49] defines prompt engineering as a technique that involves the use of task-specific hints, known as prompts, to adapt large pre-trained models to new tasks. According to Gu, "it enables the ability to perform predictions based solely on prompts without updating model parameters". Generative AI excels in understanding the English language and it is more than capable of addressing basic queries. However, refining it to perform highly specific tasks consistently requires meticulous adjustments. Specific yet concise instructions are needed to reduce AI hallucinations. Even altering a single word within a prompt can significantly impact the AI's output. This is why prompt engineering is at the heart of building applications with OpenAI and this is the new way of "fine-tuning" the AI without having to retrain the model. The chatbot's evaluation will involve rigorous testing with varying prompts.

F. Chainlit's Human Feedback Feature

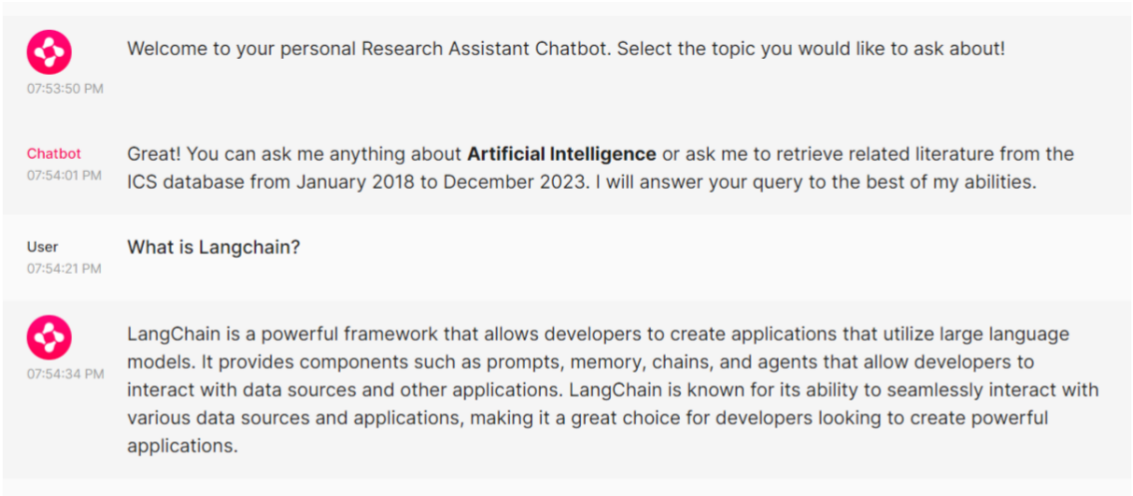
Chainlit has created a unique feature in which one can continuously give feedback to the application. It is a simple element that "allows the users to provide direct feedback on the interaction, which can be used to improve the performance and accuracy of your system." This feedback system requires users to ask queries to the chatbot and rate each answer they receive. The feedback gets stored and can be created into its own dataset that the developer can use to feed into a new iteration of the agent's knowledgebase. The feedback scores serve as a means for objective measurements and comparisons between agent versions. Accumulating this data establishes a dataset that can be crucial in improving the system's accuracy in its results. Human feedback directly facilitates model training and fine-tuning for specific interactions, which allows for a method to improve performance over time. [50]

APPENDIX I
THEORETICAL FRAMEWORK

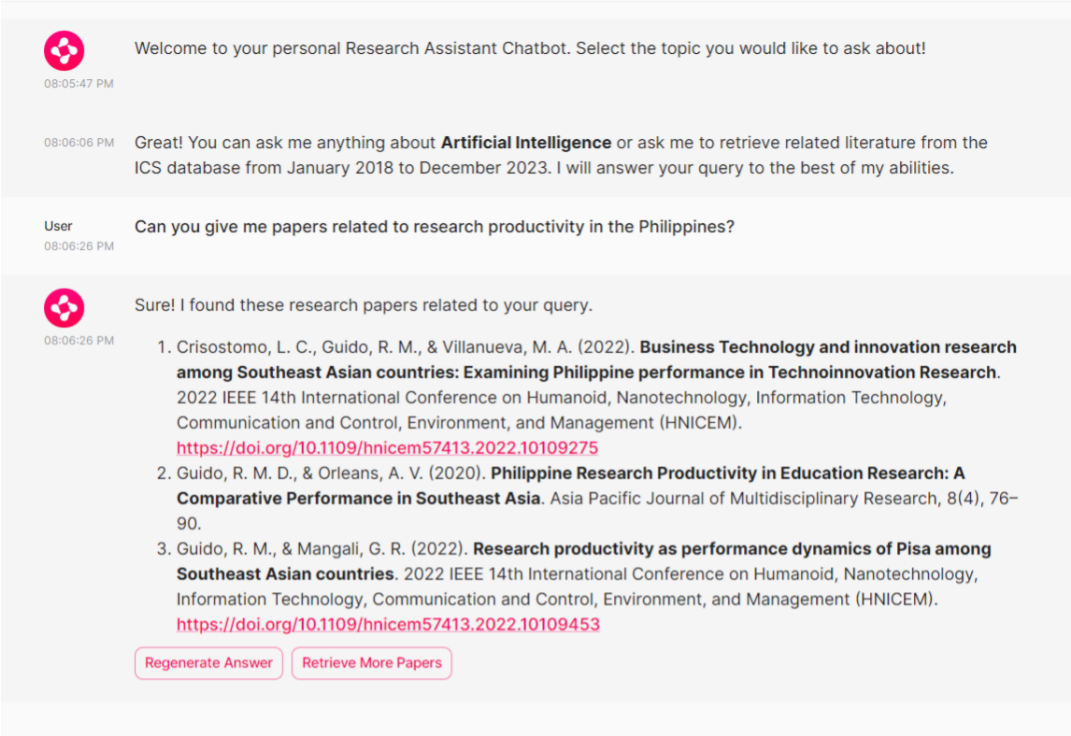


APPENDIX II
TYPE OF INQUIRY

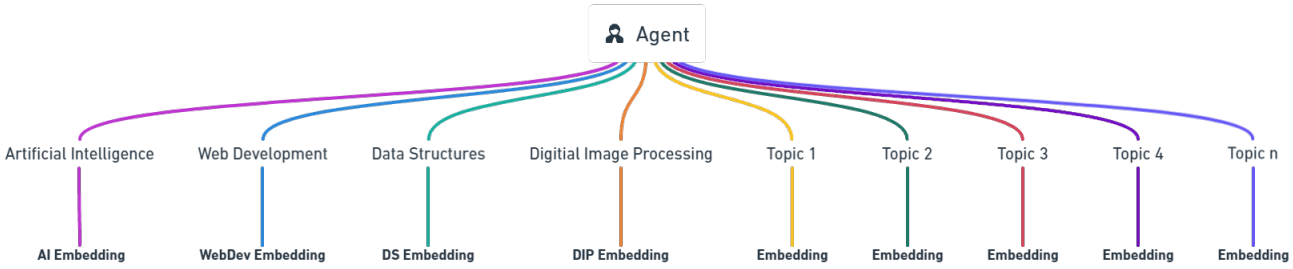
A. Screenshot of Query with a Direct Answer



B. Screenshot of a List-based Answer



APPENDIX III
AGENT FRAMEWORK



REFERENCES

[1] A. Xie, "The effect of reading in shaping undergraduates' tolerance to ambiguity," *International Journal of Language, Literature and Linguistics*, vol. 5, no. 2, pp. 51–56, Jun 2019. [Online]. Available: http://www.ijlll.org/index.php?m=content&c=index&a=show&catid=59&id=515

[2] L. Katan and C. A. Baarts, "Inquiry-based reading – towards a conception of reading as a research method," *SageJournals*, vol. 19, no. 1, pp. 58–75, 2018. [Online]. Available: https://journals.sagepub.com/doi/10.1177/1474022218760261

[3] J. T. Liwag, "Does the philippines value scientific research?" 2020, accessed: Nov. 1, 2023. [Online]. Available: https://www.cnnphilippines.com/life/culture/2020/10/21/scientific-research-philippines-barriers.html

[4] D. C. Bueno, "Research and publish or hibernate: Analysis of the limiting factors towards scientific productivity among professorial lecturers in the philippines," *CC The Journal: A Multidisciplinary Research Review*, vol. 14, p. Page range, October 2019. [Online]. Available: https://www.academia.edu/40029215/Research_and_Publish_or_Hibernate_Analysis_of_the_Limiting_Factors_towards_Scientific_Productivity_among_Professorial_Lecturers_in_the_Philippines

[5] P. Suta, X. Lan, B. Wu, P. Mongkolnam, and J. H. Chan, "An overview of machine learning in chatbots," *International Journal of Mechanical Engineering and Robotics*, vol. 9, no. 4, pp. 502–510, April 2020. [Online]. Available: https://pdfs.semanticscholar.org/7038/fecc82293642d64563cd73b04298073dbac6.pdf

[6] B. A. Shawar and E. Atwell, "Chatbots: Are they really useful?" *Journal for Language Technology and Computational Linguistics*, vol. 22, no. 1, pp. 29–49, July 2007. [Online]. Available: https://www.semanticscholar.org/paper/Chatbots%3A-Are-they-Really-Useful-Shawar-Atwell/8d8284bfba7ebcb4e2575d864ec7c16ea6a168f0

[7] C. Kooli, "Chatbots in education and research: A critical examination of ethical implications and solutions," *Sustainability*, vol. 15, no. 7, p. 5614, 2023. [Online]. Available: https://www.mdpi.com/2071-1050/15/7/5614

[8] R. Agarwal and M. Wadhwa, "Review of state-of-the-art design techniques for chatbots," *SN Computer Science*, vol. 1, no. 246, July 2020. [Online]. Available: https://link.springer.com/article/10.1007/s42979-020-00255-3

[9] OpenAI, "Introducing chatgpt," 2022. [Online]. Available: https://openai.com/blog/chatgpt

[10] T. Alqahtani, H. Badreldin, M. Alrashed, A. Alshaya, S. Alghamdi, K. bin Saleh, S. Alowais, O. Alshaya, I. Rahman, M. Al Yami, and A. Albekairy, "The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research," *Research in Social and Administrative Pharmacy*, vol. 19, no. 8, pp. 1236–1242, August 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1551741123002802?casa_token=I5hiMfxgWKQAAAAA:k2tlWqIGGs2mgovK_wcoTj9ZVqjp19mRoGf9guxIVO_3TJfst3VjGQtO9iq7HayHLzGnbgRoB5M

[11] GitHub, "Github copilot," 2021. [Online]. Available: https://github.com/features/copilot

[12] O. Topsakal and T. C. Akinci, "Creating large language model applications utilizing langchain: A primer on developing llm apps fast," *International Conference on Applied Engineering and Natural Sciences*, vol. 1, no. 1, pp. 1050–1056, July 2023. [Online]. Available: https://www.researchgate.net/publication/372669736_Creating_Large_Language_Model_Applications_Utilizing_LangChain_A_Primer_on_Developing_LLM_Apps_Fast

[13] R. M. Guido and G. Mangali, "Research productivity as performance dynamics of pisa among southeast asian countries," *Asia Pacific Journal of Multidisciplinary Research*, vol. 8, no. 4, pp. 76–90, November 2020. [Online]. Available: https://ieeexplore.ieee.org/document/10109453

[14] L. Crisostomo, R. M. Guido, and M. Villanueva, "Business technology and innovation research among southeast asian countries: Examining philippine performance in technoinnovation research," *International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management*, pp. 1–6, December 2022. [Online]. Available: https://ieeexplore.ieee.org/document/10109275?denied=

[15] S. Wa Mbaleka, "Factors leading to limited faculty publications in philippine higher education institutions," *International Forum*, vol. 18, no. 2, pp. 121–141, November 2015. [Online]. Available: https://www.researchgate.net/publication/290395664_Factors_Leading_to_Limited_Faculty_Publications_in_Philippine_Higher_Education_Institutions

[16] J. Lobo, ""behind research excellence": Students' challenges and barriers in the completion of an undergraduate thesis in the case of a premier local college in pampanga, philippines," *International Journal of Disabilities Sports and Health Sciences*, vol. 6, no. 1, pp. 38–52, February 2023. [Online]. Available: https://dergipark.org.tr/en/pub/ijdshs/issue/75824/1230630

[17] U. I. of Statistics, "Country report: Philippines," 2015. [Online]. Available: https://uis.unesco.org/en/country/ph?theme=science-technology-and-innovation

[18] J. Lim, "The philippines' scientific research is lagging, and it's due to a lack of government support," 2023, accessed: Nov. 14, 2023. [Online]. Available: https://www.cnnphilippines.com/life/culture/2020/10/21/scientific-research-philippines-barriers.html

[19] E. R. Deyro, "How a team of filipino scientists developed a covid-19 test kit," March 2020, accessed: Nov. 14, 2023. [Online]. Available: https://www.cnnphilippines.com/life/culture/2020/3/13/covid-test-kit-scientists.html?fbclid

[20] A. Index, "Scientific index 2024: Rankings for scientists in philippines," 2024, accessed: Dec. 2, 2023. [Online]. Available: https://www.adscientificindex.com/?country_code=ph

[21] J. Cornelio, "The state of research in the philippines," November 2023, accessed: Nov. 23, 2023. [Online]. Available: https://www.rappler.com/voices/thought-leaders/opinion-state-of-research-philippines/

[22] SciVal, "Scival: Comprehensive research analytics solution," 2023. [Online]. Available: https://www.scival.com/home

[23] ChatPDF, "Chatpdf: Your pdf ai," 2023.

[24] Elicit, "Elicit: Ai solutions," 2023. [Online]. Available: https://elicit.com/

[25] A. Agrahari, H. Shaikh, A. Pal, A. L. Yadav, and A. Singhal, "A survey of various chatbot implementation techniques," *International Journal of Research in Engineering, Science and Management*, vol. 1, no. 12, December 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:212484172

[26] E. Adamopoulou and L. Moussiades, "Chatbots: History, technology, and applications," *Machine Learning with Applications*, vol. 2, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827020300062

[27] B. Borah, D. Pathak, P. Sarmah, B. Som, and S. Nandi, *Survey of Textbased Chatbot in Perspective of Recent Technologies*, June 2019, pp. 84–96.

[28] A. Haque, "A brief analysis of "chatgpt" – a revolutionary tool designed by openai," *EAI Endorsed Transactions on AI and Robotics*, vol. 1, no. 1, March 2023. [Online]. Available: https://pdfs.semanticscholar.org/5e59/9b60cf4bbe22bb8254dc33a80dde0b56241c.pdf

[29] L. Fan, L. Li, Z. Ma, S. Lee, H. Yu, and L. Hemphill, "A bibliometric review of large language models research from 2017 to 2023," April 2023. [Online]. Available: https://arxiv.org/abs/2304.02020

[30] MindsDB, "Navigating the llm landscape: A comparative analysis of leading large language models," 2022 2023. [Online]. Available: https://mindsdb.com/blog/navigating-the-llm-landscape-a-comparative-analysis-of-leading-large-...language-models

[31] S. Abdelhamid, "Using chatbots as smart teaching assistants for first-year engineering students," July 2020. [Online]. Available: https://www.semanticscholar.org/paper/Using-Chatbots-as-Smart-Teaching-Assistants-for-Abdelhamid/2a0079fba5001991701c770979480fb854f3c358

[32] A. Folstad, T. Araujo, E. L.-C. Law, P. B. Brandtzaeg, S. Papadopoulos, L. Reis, M. Baez, G. Laban, P. McAllister, C. Ischen, R. Wald, F. Catania, R. M. von Wolff, S. Hobert, and E. Luger, "Future directions for chatbot research: an interdisciplinary research agenda," *Computing*, vol. 103, p. 2915–2942, October 2021. [Online]. Available: https://link.springer.com/article/10.1007/s00607-021-01016-7#citeas

[33] B. Klimova and P. M. I. Seraj, "The use of chatbots in university efl settings: Research trends and pedagogical implications," *Frontiers in Psychology*, vol. 14, March 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1131506/full

[34] B. Memarian and T. Doleck, "Chatgpt in education: Methods, potentials, and limitations," *Computers in Human Behavior: Artificial Humans*, vol. 1, no. 2, December 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2949882123000221

[35] B. Lund and T. Wang, "Chatting about chatgpt: How may ai and gpt impact academia and libraries?" *Library Hi Tech News*, January 2023. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4333415

[36] A. Sreeram and P. J. Sai, "An effective query system using llms and langchain," *International Journal of Engineering Research and*

Technology (IJERT), vol. 12, no. 6, June 2023. [Online]. Available: https://www.ijert.org/an-effective-query-system-using-llms-and-langchain
[37] LangChain, "Langchain documentation - introduction," 2023. [Online]. Available: https://python.langchain.com/docs/get_started/introduction
[38] A. M. Rouhi, "Scientists defy dire conditions: Researchers in the philippines use various strategies to keep some good work going," *Chemical and Engineering News*, vol. 74, pp. 57–61, November 1996. [Online]. Available: https://api.semanticscholar.org/CorpusID:155243704
[39] K. I. Navarro and M. McKinnon, "Challenges of communicating science: perspectives from the philippines," *Environmental Science, Political Science, Sociology*, February 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:211551375
[40] A. Calma, "Challenges in preparing academic staff for research training and supervision : The case of the philippines," *International Journal of Educational Management*, vol. 28, pp. 705–715, August 2014. [Online]. Available: https://api.semanticscholar.org/CorpusID:142518133
[41] L. M. Coloquit, "Philippine education in the modern world: A trench for global academic success or another year of educational failure?" *Journal of English Education and Linguistics*, January 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:253542522
[42] E. Ferrara, "Should chatgpt be biased? challenges and risks of bias in large language models," April 2023. [Online]. Available: https://arxiv.org/abs/2304.03738
[43] D. Cirillo, S. Catuara-Solarz, C. Morey, E. Guney, L. Subirats, S. Mellino, A. Gigante, A. Valencia, M. J. Rementeria, A. S. Chadha, and N. Mavridis, "Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare," *Digital Medicine*, vol. 3, no. 81, June 2020. [Online]. Available: https://www.nature.com/articles/s41746-020-0288-5
[44] M. Sallam, "The utility of chatgpt as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations," *Healthcare*, February 2023. [Online]. Available: https://www.medrxiv.org/content/10.1101/2023.02.19.23286155v1
[45] F. Kitamura, "Chatgpt is shaping the future of medical writing but still requires human judgment," *Radiological Society of North America*, February 2023. [Online]. Available: https://pubs.rsna.org/doi/abs/10.1148/radiol.230171?journalCode=radiology
[46] W. Hariri, "Unlocking the potential of chatgpt: A comprehensive exploration of its applications, advantages, limitations, and future directions in natural language processing," *ArXiv*, March 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:257952074
[47] T. Lopez and M. Qamber, "The benefits and drawbacks of implementing chatbots in higher education," March 2022. [Online]. Available: https://www.diva-portal.org/smash/get/diva2:1673913/FULLTEXT01.pdf
[48] M. Salvagno, F. S. Taccone, and A. G. Gerli, "Can artificial intelligence help for scientific writing?" *Critical Care*, vol. 27, no. 75, February 2023. [Online]. Available: https://ccforum.biomedcentral.com/articles/10.1186/s13054-023-04380-2
[49] J. Gu, Z. Han, S. Chen, A. Beirami, B. He, G. Zhang, R. Liao, Y. Qin, V. Tresp, and P. Torr, "A systematic survey of prompt engineering on vision-language foundation models," *ArXiv*, July 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:260357841
[50] Chainlit, "Chainlit overview," 2023. [Online]. Available: https://docs.chainlit.io/get-started/overview

Louise Gabrielle L. Talip Loys is a senior BS Computer Science student studying at the University of the Philippines Los Banos. She is a member of the Alliance of Computer Science Students UPLB and the Parliament: UPLB Debate Society. She has also been an member of the ICS competitive programming club, UPLB Eliens, and the ICS CTF club, CLK_TCK. Her adviser is Prof. Reginald Neil C. Recario. She has constantly been a part of the Top 50 CAS Students and was awarded as the CAS Outstanding Student in the Sciences in 2021. She primarily focuses on research related to Artificial Intelligence, specifically in Generative AI, and she has experience in using technological stacks for web development.