AI-enabled Intelligent Assistant to Improve Reading and Comprehension Skills in English Language

RP 24-25J-027

Project Proposal Report

S. A. D. S. Kumarathunga IT21118340

B.Sc. (Hons) in Information Technology Specializing in Software Engineering

Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

August 2024

AI-enabled Intelligent Assistant to Improve Reading and Comprehension Skills in English Language

RP 24-25J-027

Project Proposal Report

S. A. D. S. Kumarathunga IT21118340

B.Sc. (Hons) in Information Technology Specializing in Software Engineering

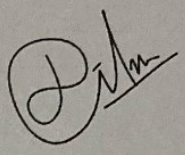
Department of Computer Science and Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

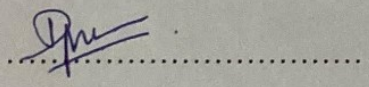
August 2024

DECLARATION

We declare that this report represents our original work. It does not incorporate any material from previous submissions for degrees or diplomas at other universities or institutes of higher learning without proper acknowledgment. To the best of our knowledge and belief, this proposal does not contain any previously published or written material by others, unless such material is properly acknowledged within the text.

Name	Student Number	Signature
S. A. D. S. Kumarathunga	IT21118340	

The above candidate is carrying out research for the undergraduate Dissertation under supervision of the undersigned.



.....

Signature of the Supervisor

26/8/2024

.....

Date

ACKNOWLEDGEMENT

I am deeply grateful to my dissertation supervisors, Professor Dasuni Nawainna and Mr. Jeewaka Perera from the Faculty of Computing, whose insightful guidance and unwavering support were instrumental throughout my research journey. Their astute observations and constructive critiques played a pivotal role in shaping my research and elevating its overall quality.

I extend my heartfelt thanks to the CDAP team for their invaluable assistance and encouragement throughout my academic pursuits. Their commitment to both teaching and research has been a constant source of inspiration for me.

I would also like to express my sincere appreciation to our esteemed panel members for their thoughtful comments and feedback throughout the research process. Their contributions have significantly enhanced the depth and validity of my work.

Lastly, I wish to acknowledge the invaluable support and collaborative spirit of my fellow team members. Our engaging discussions and idea-sharing sessions have greatly enriched my understanding of the subject matter and broadened my perspectives.

ABSTRACT

This research explores the application of Artificial Intelligence (AI) in enhancing reading and comprehension skills in the English language, focusing on the development of AI-enabled platforms capable of dynamically generating adaptable content and providing automated assessments and feedback for essay-type questions. By integrating Large Language Models (LLMs) and prompt engineering techniques, this platform aims to replicate human-like evaluations, offering personalized learning experiences tailored to individual learners' needs. The study investigates the effectiveness of LLM in understanding and grading student answers and the utility of LLMs in creating adaptive content. Through a comprehensive review of recent research, this paper highlights the current state and future directions of AI applications in education, underscoring the transformative potential of AI in improving English language proficiency and fostering a more equitable educational landscape.

Keywords: Artificial Intelligence (AI), Large Language Models (LLMs), Prompt Engineering

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

- Gen AI - Generative AI
- LLM - Large Language Model
- RAG - Retrieval-Augmented Generation

1. INTRODUCTION

1.1 Background

Advanced comprehension skills are crucial for mastering the English language. These skills include the ability to understand the implied meaning of a text, evaluate and analyze the content, and make inferences and deductions based on the information provided. Essential components of these skills involve connecting different parts of the text, applying background knowledge, identifying the main idea, finding important facts and supporting details, summarizing content, and generating as well as asking relevant questions.

The importance of reading comprehension skills in English language learning cannot be overstated, especially in an era where information is readily available yet requires critical evaluation. English, as a global language, is a gateway to vast amounts of information, and proficiency in comprehension is key to accessing, processing, and applying this information effectively. However, despite its importance, many learners struggle with advanced comprehension tasks due to factors such as limited vocabulary, lack of engagement, and insufficient practice in higher-order thinking skills.

This challenge has sparked significant interest in leveraging Artificial Intelligence (AI) to improve reading and comprehension skills. AI-enabled platforms offer personalized learning experiences, adapting to individual learner needs and providing real-time feedback. These platforms can analyze learner behavior, identify weaknesses, and offer targeted exercises to strengthen comprehension skills. The effectiveness of AI in educational contexts has been widely studied, but its application to enhance reading comprehension in English, particularly in non-native speakers, remains an area ripe for exploration.

1.2 Literature Survey

The integration of AI in education, particularly for enhancing reading and comprehension skills in English, has gained considerable traction in recent years. This review explores the current state of research in this area, focusing on AI-enabled platforms designed to dynamically generate adaptable content, automatically assess essay-type responses, and provide feedback. By analyzing key studies in this domain, we aim to contextualize the advancements and challenges in utilizing AI for educational purposes, with a specific emphasis on improving English language proficiency.

Lack of Dynamically Adaptable Content

A key challenge in the realm of AI-enabled educational tools is the lack of dynamically adaptable content that can respond to individual learners' needs. While several studies have explored the application of AI and machine learning in educational contexts, the focus has often been on static content delivery rather than content that adapts in real-time to a student's comprehension level.

Mirabal et al. (2023) [2] examine the use of deep learning-based language models to improve reading comprehension. Their study demonstrates that language models can enhance reading comprehension by offering personalized recommendations. However, the adaptability of content is limited to predefined pathways and lacks real-time dynamic adaptation. The approach does not fully address the individual learner's evolving needs during a reading session. This shortcoming is significant, as students' comprehension levels can vary widely, necessitating a more responsive system that adjusts content complexity and format on the fly.

Laban et al. (2022) [3] delve into automated question generation for quiz design, aiding teachers in creating quizzes that align with the material covered. This research is commendable in that it provides a means of customizing assessments, yet it falls short in dynamically adapting content based on real-time analysis of a student's responses. The quiz design task remains predominantly teacher-driven, and while it supports personalized assessment to some extent, it does not leverage AI's full potential to modify instructional content dynamically during the learning process.

In contrast, the proposed research aims to develop an AI-enabled platform that can provide real-time, dynamically adaptable content tailored to individual learners. By employing advanced natural language processing (NLP) techniques and adaptive learning algorithms, this platform would continuously assess the learner's comprehension and adjust the reading materials accordingly. This approach would address the identified gaps in existing research by moving beyond static content delivery to a more fluid, responsive educational experience.

Lack of Tools for Automated Essay-Type Question Evaluation

The second major challenge is the absence of reliable tools for the automated assessment of essay-type questions, which traditionally require human evaluation due to the complexity and subjectivity involved in grading. This area has seen considerable research, yet existing solutions remain inadequate for high-stakes educational environments.

Moholkar et al. (2024) [1] provide a comprehensive survey of machine learning techniques for evaluating descriptive answers. Their survey covers various approaches, including supervised learning models and ensemble methods, highlighting their effectiveness in grading to some extent. However, the authors acknowledge the limitations in these models' ability to handle the nuances of human language, such as context, tone, and rhetorical devices, which are crucial for evaluating essay-type questions. The models surveyed often struggle with maintaining consistency and fairness in grading, especially in cases requiring deep contextual understanding.

Similarly, Xia et al. (2024) [4] conduct an empirical study on large language models (LLMs) as automated essay scoring tools, using the TOEFL Independent Writing Task as a case study. Their findings suggest that while LLMs can perform reasonably well in scoring essays, there are significant drawbacks, including the models' susceptibility to being misled by superficial features like essay length or vocabulary sophistication. Furthermore, the lack of transparency in LLM decision-making processes raises concerns about their reliability and fairness in educational assessments.

Hussein et al. (2019) [5] review automated language essay scoring systems, identifying a range of methodologies from traditional rule-based systems to more contemporary machine learning approaches. While this review highlights progress in the field, it also points out the persistent challenges in achieving human-like grading accuracy, especially in handling creative or argumentative essays that require nuanced judgment.

The proposed research intends to overcome these limitations by developing a novel assessment tool that combines the strengths of machine learning with advanced linguistic analysis to better emulate human evaluators. This tool will focus on understanding the content and structure of essays more deeply, including context, argumentation, and rhetorical devices, thereby improving grading accuracy and fairness. Moreover, by integrating this tool with the dynamically adaptable content system, the platform would not only assess but also provide immediate, actionable feedback, thereby enhancing the learning process in real-time.

Conclusion

The research landscape for AI-enabled platforms aimed at improving reading and comprehension skills in English is marked by significant advancements but also by critical gaps, particularly in the areas of dynamically adaptable content and automated essay-type question evaluation. Existing studies have laid important groundwork but have not fully addressed these challenges. The proposed research seeks to fill these gaps by developing a more responsive and nuanced AI-enabled platform that provides both dynamic content adaptation and robust assessment tools for essay-type questions. By doing so, it aims to offer a more personalized and effective learning experience, bridging the current shortcomings in educational technology.

1.3 Research Gap

The research on AI-enabled platforms to improve reading and comprehension skills in the English language identifies significant gaps in the existing technological landscape, particularly concerning the lack of dynamically adaptable content and tools to assess and provide feedback for essay-type questions. Current educational technologies are generally limited in their capacity to adjust content dynamically based on individual learner needs. This gap is particularly detrimental in the context of English language education, where learners often require personalized instruction tailored to their specific linguistic challenges and progress levels. Although deep learning models have shown potential in generating human-like text and adaptive learning pathways, they remain constrained by their inability to continuously modify content in real-time based on the user's evolving comprehension skills[6]. Existing research highlights that most platforms rely on static content, which fails to engage learners or address their individual learning trajectories effectively [7]. This lack of adaptability hinders the development of nuanced reading and comprehension skills, as learners are not provided with content that is both challenging and accessible.

Moreover, the absence of robust tools to evaluate essay-type questions presents a significant barrier to developing critical reading and writing skills. Essays require not only a deep understanding of content but also the ability to articulate thoughts coherently and persuasively. Despite advancements in natural language processing (NLP), current AI systems struggle to provide accurate, context-sensitive feedback on essays, which is essential for nurturing higher-order cognitive skills [8]. The complexity of evaluating argumentation, creativity, and coherence in essays is beyond the capabilities of most automated systems, necessitating human intervention and thus limiting scalability. In English language learning, where essay writing is crucial for assessing comprehension and language proficiency, this gap is particularly problematic. Therefore, there is a pressing need for AI-enabled platforms that can not only adapt content dynamically but also assess and provide meaningful feedback on essay-type questions, bridging the gap between personalized learning and scalable, high-quality education.

Comparison with Similar Products:

Apps	Dynamic Content	Evaluate answers without human involvement	Provide feedback without human involvement
Generic Reading Apps	✗	✗	✗
Duolingo	✓	Only offer MCQ based questions	Limited feedback
AceReader Pro	✗	✗	✗
ReadyRead	✗	✗	✗
Smart AI Reading Assistant for Reading Comprehension	✗	✗	✗
Proposed System	✓	✓	✓

Table 1.1 Comparison of Similar Products

When compared to existing products, the limitations of current reading and comprehension apps become more apparent. Generic reading apps, for instance, typically offer a limited amount of static content, which does not adapt to the learner's progress or interests. This static nature fails to engage users over time and does not cater to the individual learning paths necessary for effective reading comprehension.

AceReader Pro[9], while aiming to improve reading comprehension through fast reading techniques, focuses predominantly on speed, which may not be suitable for all students, especially those still developing basic reading skills. Furthermore, AceReader Pro's[9] reliance on pre-existing content limits its adaptability and fails to align with the user's personal interests or learning needs.

Similarly, ReadyRead[10], a mobile application designed as supplementary material for reading comprehension, primarily addresses the issue of material insufficiency in the classroom but suffers from similar limitations. The app relies heavily on static, pre-created content and lacks personalized learning paths, thereby failing to provide dynamic content needed for effective skill development.

Most critically, none of these products have the capability to assess or provide feedback on essay-type questions, a crucial aspect of language learning that typically requires human evaluation. This gap underscores the need for a more sophisticated, AI-driven platform that can deliver both dynamic content and nuanced assessment tools, addressing the comprehensive needs of English language learners.

1.4 Research Problem

1. Lack Of Dynamically Adaptable Content

In contemporary English education, one of the most pressing challenges is the lack of dynamically adaptable content that can adjust to the varying needs and interests of learners. This contributes to student dissatisfaction, as they are presented with limited options for reading materials and content. As a result, students often find themselves bored and less engaged, which significantly hampers their ability to achieve high-level comprehension skills.

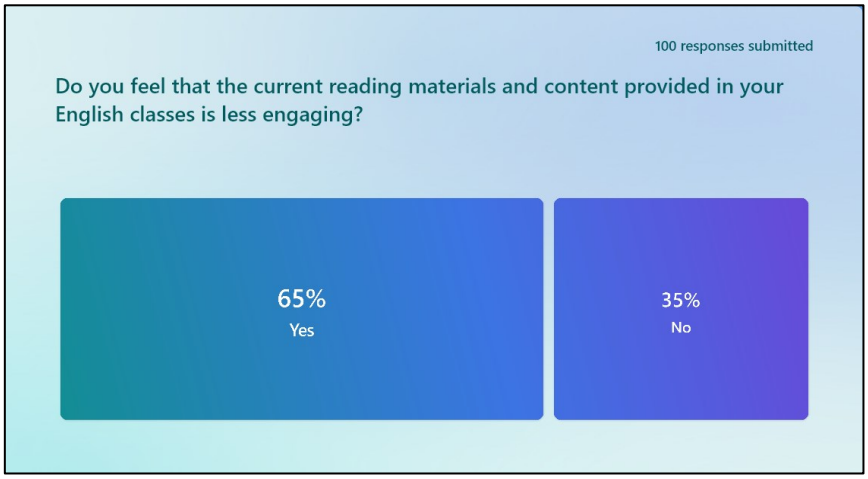


Figure 1.1 Survey report on student opinion regarding static reading materials

Introducing reading materials and exercises that dynamically adapt to student needs and interests can markedly improve student engagement with the materials, thereby enhancing comprehension skills. Such adaptability ensures that the content remains relevant and challenging, keeping students motivated and invested in their learning journey.

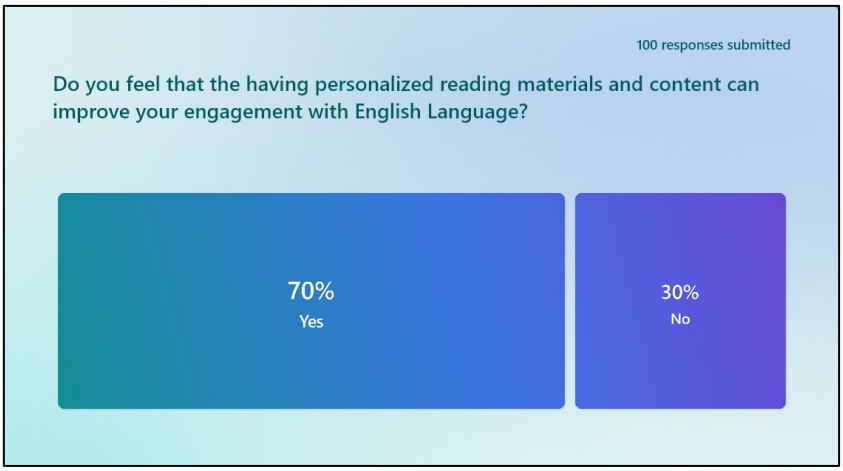


Figure 1.2 Survey report on student opinion regarding personalized reading materials

2. Lack tools to provide assessments for essay-type questions that typically require human evaluation.

Another significant challenge in English education is the absence of tools capable of autonomously evaluating answers for essay-type questions. This limitation forces students to rely on their English teachers for feedback, a process that is not only time-consuming but also inefficient. Students often have to wait for extended periods to receive feedback on their work, which not only wastes valuable time but also limits the opportunities for immediate correction and skill enhancement. The delay in receiving feedback can hinder the development of critical thinking and writing skills, which are essential for success in English education and beyond.

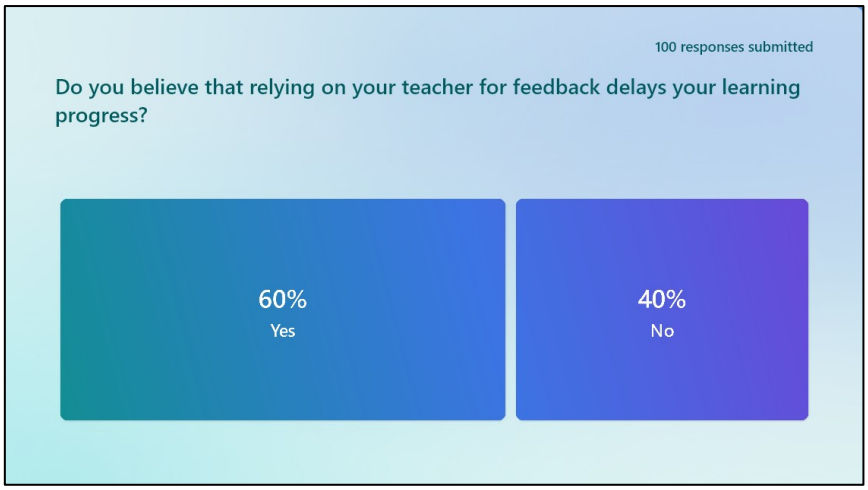


Figure 1.3 Survey report on student opinion regarding manual answer evaluation process

Having tools that automatically provide assessments and feedback for essay-type questions can significantly improve the student learning process. Such tools can offer immediate, personalized feedback, allowing students to identify and rectify errors promptly. This immediacy in feedback is crucial for facilitating rapid learning cycles and encouraging students to actively engage with their work, thereby enhancing their comprehension skills and overall academic performance.

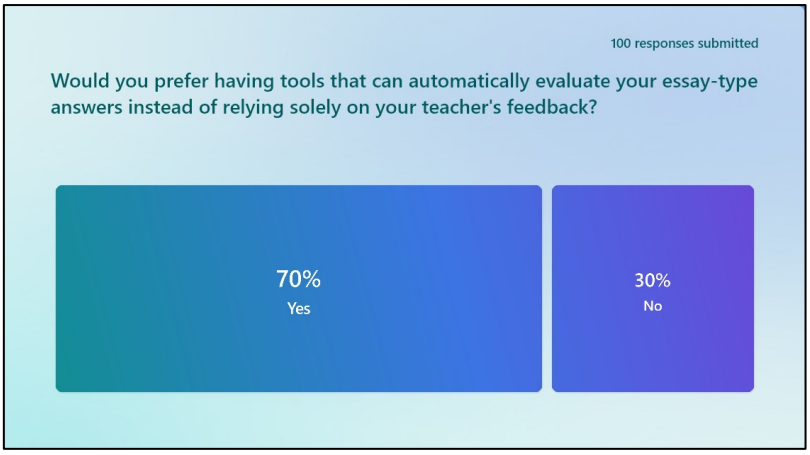


Figure 1.4 Survey report on student opinion regarding automated answer evaluation process

By addressing these challenges, educators can create a more dynamic, efficient, and personalized learning environment that better serves the needs of English language learners, paving the way for improved educational outcomes.

2. RESEARCH OBJECTIVES

2.1 Main Objective

The main objective of this study is to develop a web application through which English language learners can improve advanced comprehension skills, focusing on areas such as sequencing, summarizing, and self-questioning.

This initiative aims to address the growing need for effective educational tools that cater to the complexities of the English language, particularly in enhancing learners' abilities to analyze, interpret, and engage critically with texts.

2.2 Specific Objectives

There are two specific objectives that must be reached in order to achieve the overall objective described above.

1. **Utilizing an LLM model to generate dynamic content, including lessons and exercises tailored to the user's interests**

This involves developing effective prompts that encourage the model to produce relevant and engaging material. Additionally, determining the optimal method to integrate RAG with the LLM is crucial for enhancing the content's accuracy and depth. By achieving these objectives, the application will offer personalized learning experiences that cater to individual user preferences, significantly improving engagement and comprehension.

2. **Leverage the LLM model to accurately evaluate user answers and provide feedback based on given criteria without human involvement**

This requires careful prompt engineering to ensure the model understands the task at hand and delivers constructive feedback. Integrating RAG with the LLM will further refine this process, allowing for more precise evaluations and comprehensive feedback mechanisms. This functionality will automate the grading process, saving time and resources while enhancing the learning experience by offering immediate insights into performance and areas for improvement.

3. METHODOLOGY

3.1 Research Domain

3.1.1 Generative AI

Generative AI refers to a subset of artificial intelligence technologies that specialize in creating new and original content, ranging from text and images to complex synthetic data and even deepfakes. This technology leverages advanced neural network techniques such as transformers, generative adversarial networks (GANs), and variational autoencoders (VAEs) to autonomously generate a wide array of outputs. Unlike traditional AI algorithms that operate based on predefined rules, generative AI excels in tasks requiring the creation of new content, especially in natural language processing (NLP) domains. It typically begins with a prompt or dataset to guide the content generation process, allowing for iterative exploration of content variations. The advent of generative AI has opened up innovative possibilities across various sectors, from creative fields to problem-solving, promising to revolutionize how we interact with technology and create content.

3.1.2 Large Language Model (LLM)

Large Language Models (LLMs) represent a significant advancement in artificial intelligence, particularly in the realm of natural language processing. These models leverage deep learning techniques and vast datasets to understand, summarize, generate, and predict new content, essentially serving as sophisticated tools for analyzing and producing text-based content. The cornerstone of LLMs lies in their ability to learn from massive amounts of text data, acquiring a deep understanding of language nuances, syntax, and semantics. This capability enables them to perform a wide range of tasks, from answering queries to generating creative content, making them invaluable in various applications across industries. The architecture of modern LLMs, often based on transformer models, allows for efficient processing and generation of large-scale text data, setting the stage for further innovations in AI-driven content creation and analysis.

3.1.3 Prompt Engineering

Prompt engineering is a critical process in the development and refinement of generative AI systems, including Large Language Models (LLMs) and other generative AI tools. It involves carefully crafting input prompts that influence the direction, quality, and relevance of the output produced by these models. By designing effective prompts, developers can steer the AI towards specific types of content, answer formats, or behaviors, optimizing the model's performance for particular tasks or applications. This technique requires a blend of linguistic expertise and creative thinking to fine-tune prompts, extracting the most desirable outcomes from the AI models.

3.1.4 Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) is a sophisticated technique that enhances the capabilities of Large Language Models (LLMs) by incorporating external knowledge sources during the generation process. This method optimizes the output of LLMs by referencing an authoritative knowledge base outside of their initial training data, ensuring that the generated responses remain relevant, accurate, and useful across various contexts. RAG works by introducing an information retrieval component that pulls relevant information from external data sources based on the user's input. This newly acquired knowledge, combined with the LLM's existing training data, enables the model to generate more informed and contextually rich responses. The integration of RAG with LLMs addresses several limitations of standalone LLMs, such as the challenge of accessing updated information and the need for domain-specific knowledge, thereby significantly improving the accuracy and reliability of AI-generated content.

3.2 How Proposed Technologies Address Identified Research Problems

The primary research problems highlighted involve the scarcity of dynamically adaptable content and the absence of tools capable of autonomously assessing and providing feedback for essay-type questions, which typically require human evaluation. To tackle these issues, the proposed technologies—Generative AI, Large Language Models (LLM), Prompt Engineering, and Retrieval-Augmented Generation (RAG)—are chosen for their complementary strengths. Generative AI and LLMs serve as foundational technologies, capable of generating and understanding complex textual content. However, their effectiveness is significantly enhanced through Prompt Engineering, which fine-tunes the inputs to these models, ensuring the generation of relevant and meaningful content. RAG further elevates the capabilities of LLMs by integrating information retrieval, allowing the models to draw upon external databases to inform their responses, thereby addressing the issue of dynamically adaptable content.

Moreover, the combination of these technologies enables the system to autonomously assess essay-type questions and provide extensive feedback, eliminating the reliance on human evaluators. The selection of these technologies reflects a thoughtful consideration of their strengths and how they can be combined to overcome the identified research problems. This solution promises to revolutionize educational assessment, making it more accessible, personalized, and effective, thereby addressing a significant gap in current educational practices.

3.3 System Architecture

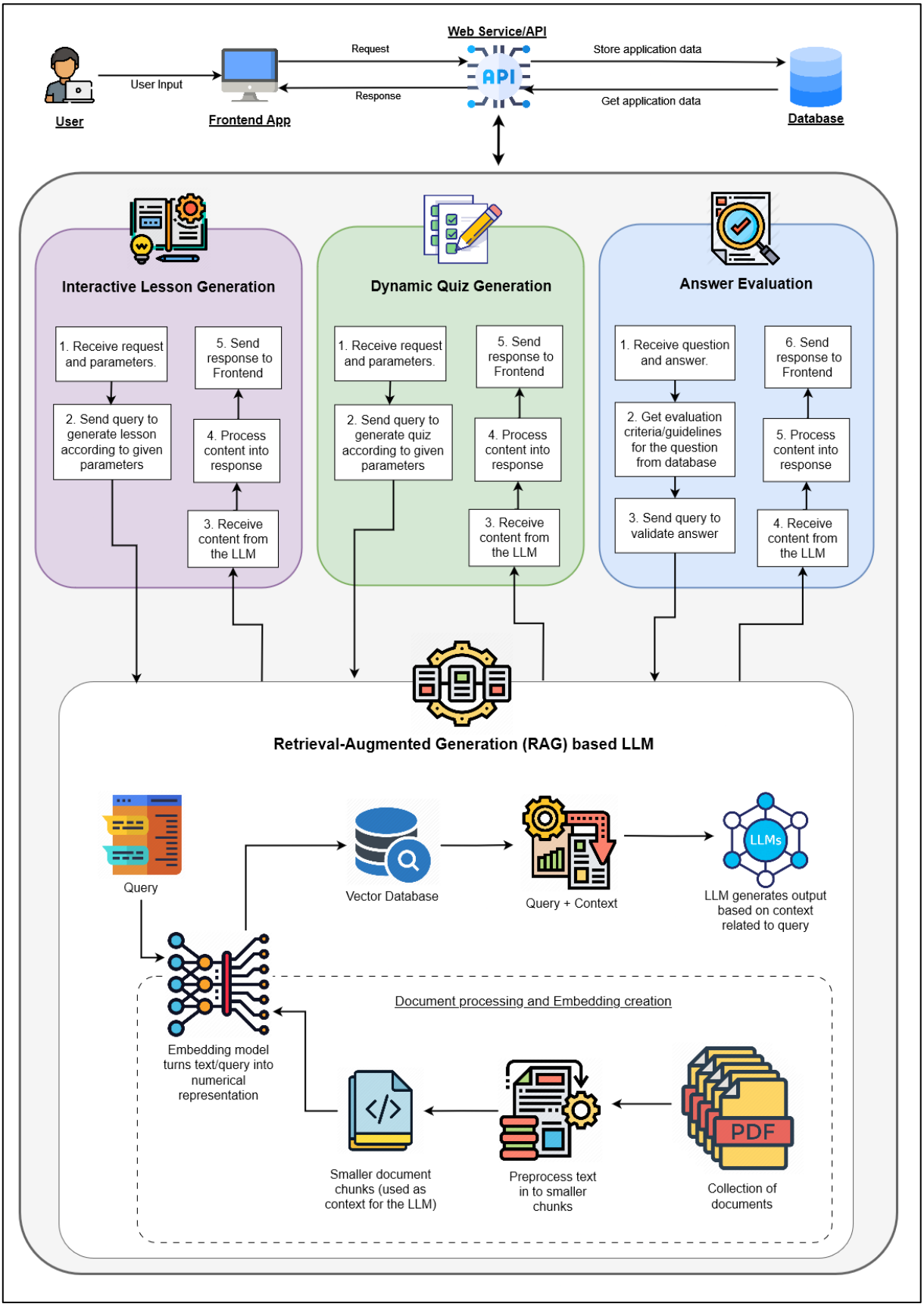


Figure 3.1 System Architecture

Above Figure 3.1 illustrates the overall high-level architectural diagram of the proposed component on Advanced Comprehension Skills Enhancement, which comprises three interconnected modules designed to enhance educational outcomes through interactive engagement and precise assessment.

The first module, Interactive Lesson Generation, employs prompt engineering and Retrieval-Augmented Generation (RAG) techniques to dynamically create content tailored to the user's interests. This approach ensures that lessons are relevant and engaging, fostering active participation and deeper comprehension.

The second module, Dynamic Quiz Generation, similarly utilizes prompt engineering and RAG to produce quizzes that adapt to the user's progress and areas of interest. This adaptive quizzing mechanism not only challenges the learner at their current level but also encourages continuous improvement by adjusting the difficulty and content of the questions in real-time.

The third and final module, Answer Evaluation, is pivotal in assessing the user's responses to the generated quizzes. Leveraging advanced algorithms, this module evaluates answers against predefined criteria without the need for human intervention, thereby automating the grading process. Furthermore, it provides extensive feedback on the user's answers, highlighting strengths and areas for improvement. This feedback loop is crucial for promoting self-reflection and learning, enabling users to understand their mistakes and learn from them effectively.

The Retrieval-Augmented Generation (RAG) pipeline involves two primary stages: Document Processing and Embedding Creation, followed by Search and Answer. In the first stage, documents are collected and preprocessed into smaller chunks, which serve as context for the Large Language Model (LLM). An embedding model then transforms these text snippets into numerical representations, storing them in a vector database. The second stage begins with a user query, which is similarly embedded. The vector database is searched to find relevant context, and the query along with its matching context is passed to the LLM. Finally, the LLM generates an output based on the context provided, creating a coherent and informative response to the original query. This hybrid approach combines the strengths of retrieval systems and generative models to produce more accurate and relevant responses.

Together, these modules form a cohesive system aimed at enhancing advanced comprehension skills through interactive learning and immediate, automated feedback. By integrating RAG-based technologies, the proposed component seeks to revolutionize educational experiences, making them more personalized, efficient, and effective.

3.4 Tools

Frontend Application	JavaScript & React
Backend API	Python & Django
Database	Postgres SQL
Vector Database	Pinecone DB
LLM Model	Llama
LLM Orchestration	Lang Chain
CI/CD	GitHub, Docker, Kubernetes
IDE	Visual Studio Code / Google Colab

Table 3.1 Tools

3.5 Technologies

- Prompt Engineering Techniques
    - Few-Shot Prompting - Limited number of examples are provided to guide an AI model towards producing specific outputs
    - Chain of Thought - Guides AI models to outline intermediate steps/reasoning leading to an answer, enhancing transparency and accuracy in complex problem-solving tasks.
- Retrieval-Augmented Generation (RAG) - Retrieves relevant information from databases and incorporates it into the model's output.
- Vector DB - Store data as dense vectors, enabling fast similarity searches between entities.
- Lang Chain - Open-source framework for building applications powered by large language models. It enables developers to create custom prompts, retrieve relevant information, and generate responses efficiently, simplifying the integration of LLMs into various projects and workflows.

3.6 Software Development Life Cycle

The Software Development Life Cycle (SDLC) is a structured process used to plan, design, develop, test, deploy, and maintain software systems. It encompasses several phases that guide the development process from conception to delivery. The main goals of SDLC are to ensure high-quality software is produced efficiently and effectively. Common SDLC models include Waterfall, Spiral, and Agile methodologies, each with its own approach to managing the development lifecycle.

3.6.1 Requirements Gathering

This step involves collecting and analyzing information from various sources to define the project's scope and objectives.

Specifically for this platform, the requirements gathering process would include:

Data gathering:

- Collecting publications from the Sri Lankan Education Department to understand curriculum standards and learning objectives
- Gathering fair-use materials from the internet to build a comprehensive content library
- Ensuring compliance with copyright laws and obtaining necessary permissions

Surveying stakeholders:

- Parents: Willingness to pay for premium services, Understanding parental expectations and concerns,
- Students: Identifying specific learning needs and preferences, Preferred features and functionalities
- Educators: Gaining insights into current teaching methods and desired outcomes

This comprehensive approach ensures that the platform addresses the needs of all stakeholders involved in English language education in Sri Lanka.

3.6.2 Feasibility Study

A feasibility study is conducted to assess whether the proposed project is viable and practical.

Economic Feasibility:

- Estimating startup costs and ongoing expenses
- Analyzing potential revenue streams (subscription fees, advertising)
- Evaluating cost-benefit ratio and return on investment

Scheduled Feasibility:

- Creating a project timeline with milestones and deadlines
- Assessing dependencies between different phases of development
- Identifying critical path activities that could delay the project

Technical Feasibility:

- Evaluating the availability of necessary technologies and tools
- Assessing the complexity of implementing RAG (Retrieval-Augmented Generation) pipelines
- Considering scalability and future-proofing the architecture

This step helps in determining whether the project is technically, economically, and practically feasible, and if so, what modifications might be needed to ensure success.

3.6.3 Design

Use Case Diagram

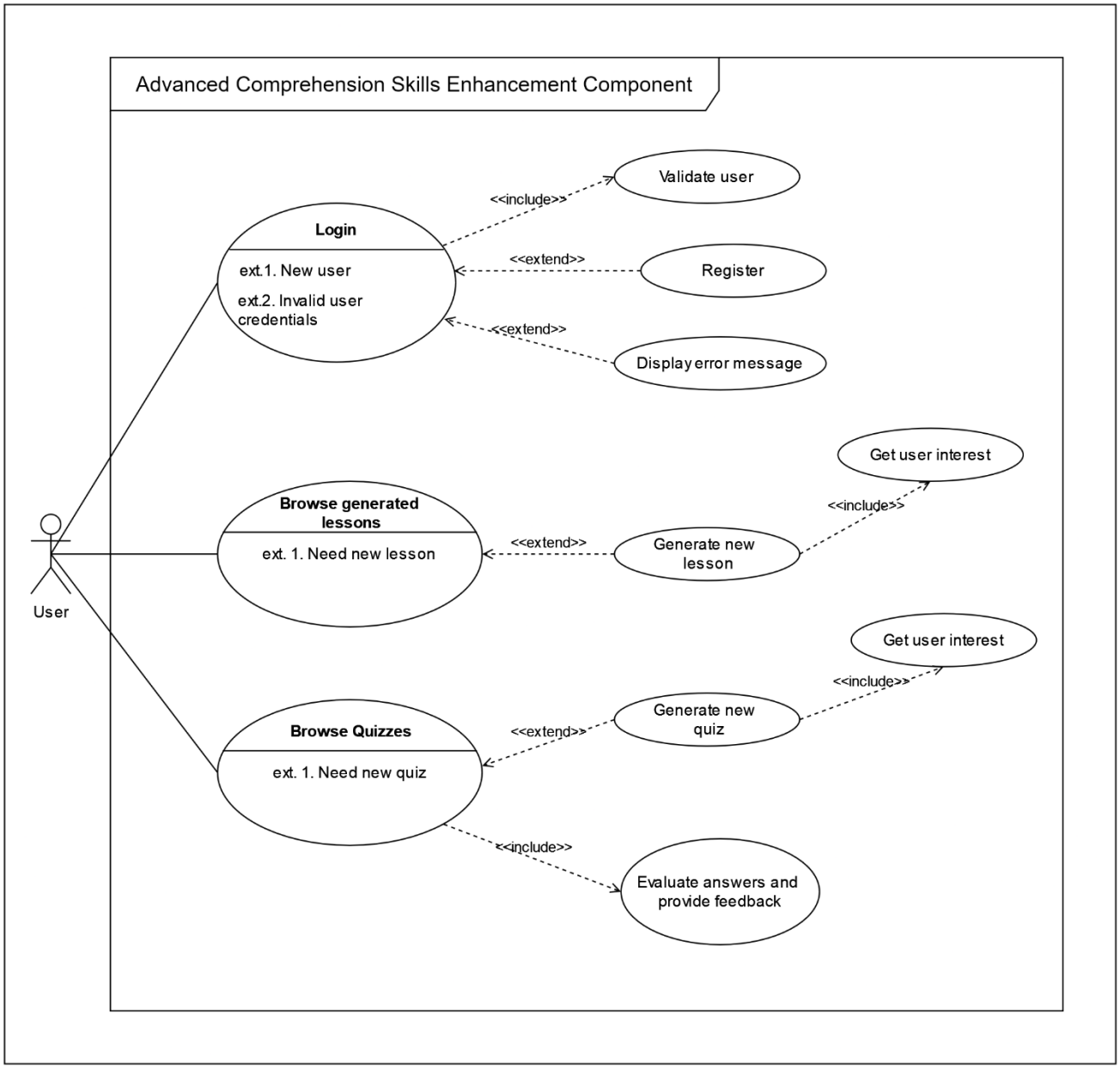


Figure 3.2 Use Case Diagram

3.6.4 Implementation

Implementation is the core phase where the actual development takes place.

Database Implementation:

- Implement a scalable database structure to store vast amounts of reading materials
- Implementing indexing strategies for efficient querying
- Ensuring data privacy and security measures

RAG Pipeline Creation:

- Developing document preprocessing techniques
- Implementing embedding models for text representation
- Creating a vector database to store and retrieve contextual information

LLM Integration:

- Selecting and integrating a suitable Large Language Model
- Integrating the RAG pipeline with the LLM
- Integrating LLM outputs into the platform's decision-making processes

Backend Development:

- Creating RESTful APIs for seamless communication between frontend and backend
- Implementing authentication and authorization systems
- Developing robust error handling and logging mechanisms

Frontend Development:

- Creating wireframes and mockups for UI design
- Implementing responsive web design principles
- Optimizing performance and user experience

3.6.5 Testing

Testing is a critical phase in the development lifecycle, ensuring that the platform meets quality standards and performs as expected.

Functional Testing:

- Verifying that all features work as designed
- Verifying user flows and interactions
- Verifying functionality of the RAG pipeline

Non-Functional Testing:

- Performance Testing: Evaluating load times and responsiveness under varying conditions
- Security Testing: Assessing vulnerabilities and implementing appropriate countermeasures
- Usability Testing: Gathering feedback on user experience and ease of use

Integration Testing:

- Verifying smooth interaction between different components (Database, LLM, Frontend)
- Ensuring data consistency across all layers of the application

System Testing:

- Comprehensive testing of the entire system in a simulated production environment
- Identifying potential bottlenecks and areas for optimization.

User Acceptance Testing (UAT):

- Conducting tests with real users to validate the platform's effectiveness
- Gathering feedback on usability, relevance, and overall value proposition

3.7 Dataset

- Publications from the Sri Lankan Education Department
- Fair-use materials from the internet

3.8 Constraints

During the development of the proposed software product, there may be constraints that limit the scope and feasibility of the research. Some of these constraints include:

- Time constraints: The research must be completed within one year, which may limit the amount of research that can be conducted.
- Limited resources: The development of the software product must stay under budget, which may limit the availability of resources such as funding, equipment, and personnel.
- Ethical considerations: The research must be conducted in an ethical manner and must comply with all relevant regulations and guidelines. Ethical clearance will be obtained from the university.
- Hardware constraints: The LLMs used by the system require significant computing power and resources to operate effectively. Limited resources may limit the availability of computing power and resources to run the models.

4. PROJECT REQUIREMENTS

4.1 Functional Requirements

- The system should be able to dynamically generate content given on user's interest and skill level.
- The system should be able to accurately evaluate user answer based on given criteria without human involvement.
- The system should be able to give extensive feedback about user's answer.

4.2 Non-Functional Requirements

- User-Friendliness - The system should provide a cross-platform application for users while maintaining an attractive, responsive and well comprehensive interphase.
- Reliability - The system should not fail or get stuck at any time throughout the process. The users should feel secure and confident while using the application. All sensitive information must be protected.
- Performance - The system should perform efficiently by providing fast and accurate results to the users.
- Availability - The application should be accessible to all users, regardless of language. It should be able to be used whenever it is needed.

4.3 System Requirements

The purpose of software requirements is to define the software resources that must be enforced on a system in order for the proposed system to function properly. The software specifications requirements for this proposed component are as follows.

- React for Frontend Application
- Django for Backend API
- Postgres SQL for Database
- Pinecone DB for Vector Database
- LLAMA for Language Model
- Lang Chain for LLM Orchestration
- VS Code to implement the code using JavaScript & Python

4.4 User Requirements

- This web application will mainly focus on learners, who are eager to develop their English language skills.

4.5 Personal Requirements

- Development team is required to gather expertise in Web frameworks, LLM related tools, Database Design, System Design & Deployment

5. WORK BREAKDOWN STRUCTURE

Phase	Tasks
Initiation Process	• Identify research problem • Topic Assessment • Projects Proposal Presentation & Report
Planning Process	• Literature Review • Requirement Analysis • Feasibility Study • Data Collection
Design Process	• Use Case Diagram • Database Design (ER diagram) • Application Architecture design • Interface Design

Implementation Process	• LLM related development
	  o Develop and Integrate RAG pipeline with LLM
	  o Develop and Test prompts for Generating content, Assessment and Provide Feedback
	• SQL Database Implementation
	• Backend API Implementation
	• Frontend App Implementation

Testing Process	• Unit Testing • Component Testing • Integration Testing • System Testing • User Acceptance Testing
Monitoring & Controlling Process	• Control Schedule & Scope
Closing Process	• Deploy Web Application • Final Report and Presentation • Publish Research Paper

Table 5.1 Work Breakdown Structure

6. GNATT CHART

Task	Duration														
	2023								2024						
	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL
1. Initial Stage															
Research topic selection															
Requirement gathering															
Study on research area															
Topic approval															
2. Proposal Stage															
Project proposal draft submission															
Proposal presentation															
3. Implementation Stage 1															
System planning and design															
Implementation of function															
Integration and testing level 1															
Progress presentation – 50%															
Prepare research paper															
4. Implementation Stage 2															
Implementation of remaining function															
Integration and testing level 2															
Progress presentation – 90%															
5. Final Stage															
Final thesis with proof reader sign off															
Deploy the website															
Final presentation															

Figure 6.1 Gnatt Chart

7. BUDGET ESTIMATIONS AND JUSTIFICATION

Expenses	Estimated Cost	Justification
Cost of cloud computing services for LLM (For training and hosting)	LKR 6000/month	The cost covers the computational resources required to train and host the Large Language Model (LLM), which is central to the app's functionality. This investment ensures the model remains up-to-date and responsive, providing accurate and timely information to users.
Cost of hosting (Frontend Application, API, Database)	LKR 7500/month	This expense supports the infrastructure necessary for the app's operation, including the frontend application, backend API, and database management. It guarantees reliable performance and security, essential for maintaining user trust and satisfaction.
Other costs (Travelling, Internet connection)	LKR 5000/month	These costs cover essential operational expenses such as traveling for meetings and events related to the app's development and promotion, as well as the internet connection needed for remote work and communication.

Table 7.1 Budget Estimations and Justifications

8. COMMERCIALIZATION

8.1 Target Audience and Market

- High school students
- English language learners
- English Educators

8.2 Business Strategy

The business model for this application centers on a Software as a Service (SaaS) model, offering two distinct tiers to cater to diverse user needs.

1. **Free Tier** supported by advertisements primarily from educational institutions, ensures broad accessibility, encouraging widespread adoption among learners. This tier serves as an entry point, introducing users to the app's features and benefits.

2. **Premium Version** is available through a subscription-based model, offering an ad-free environment with additional features and resources.

By offering these two tiers, the app can attract a diverse user base, from casual users using the free version to dedicated learners and educators opting for the premium services. This dual-tiered approach not only diversifies revenue streams but also enhances user engagement and retention, positioning the app as a valuable tool in the educational landscape.

8.3 Marketing Strategies

The marketing strategy for the app focus on establishing strategic partnerships and collaborations to enhance its reach and credibility within the educational sector. Partnering with educational institutions provides a unique opportunity to pilot the app within controlled environments, allowing for feedback and adjustments before a broader launch. These partnerships not only facilitate access to potential users but also lend legitimacy to the app, making it a trusted resource for learning.

Additionally, exploring collaborations with government agencies and non-profit organizations focused on literacy and education can further extend the app's impact. Such alliances can provide financial support, visibility, and access to networks that are crucial for scaling the product effectively. By leveraging these partnerships, the educational app can position itself as a vital tool in the educational ecosystem, reaching a wider audience and contributing significantly to improving educational outcomes.

REFERENCES

[1] K. Moholkar, M. Chaturvedi, A. Jain, A. Parkhe and K. Singh, "Machine Learning Techniques for Descriptive Answer Evaluation: A Comprehensive Survey," 2024 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2024, pp. 1412-1419, doi: 10.1109/ICICT60155.2024.10544714.

[2] P. Mirabal, M. Castillo-Sanhueza, R. Curín-Zarate and O. O. Calzadilla-Pérez, "Use of Language Models based on Deep Learning to improve reading comprehension," 2023 42nd IEEE International Conference of the Chilean Computer Science Society (SCCC), Concepcion, Chile, 2023, pp. 1-6, doi: 10.1109/SCCC59417.2023.10315757.

[3] P. Laban, C. Wu, L. Murakhovs'ka, W. Liu, & C. Xiong, "Quiz Design Task: Helping Teachers Create Quizzes with Automated Question Generation," in NAACL-HLT, 2022.

[4] W. Xia, S. Mao, & C. Zheng, "Empirical Study of Large Language Models as Automated Essay Scoring Tools in English Composition Taking TOEFL Independent Writing Task for Example," ArXiv, vol. abs/2401.03401, 2024

[5] M. A. Hussein, H. A. Hassan, & M. Nassef, "Automated language essay scoring systems: a literature review," PeerJ Computer Science, vol. 5, 2019.

[6] K. Hilali, M. Chergui and A. Ammoumou, “Adaptive Learning Systems: A Comprehensive Overview and Identification of Challenges,” 2023 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), Rabat, Morocco, 2023, pp. 1-7, doi: 10.1109/ICTMOD59086.2023.10438128.

[7] Pawar, Purushottam, “AI-Enhanced Education: Personalized Learning and Educational Technology”, 2023, doi:10.25215/9358791152.01.

[8] Wu, H., Li, S., Gao, Y. et al. “Natural language processing in educational research: The evolution of research topics”, Educ Inf Technol, 2024, doi: 10.1007/s10639-024-12764-2.

[9] Yuhua Li, "AceReader pro and reading comprehension," Proceeding of the International Conference on e-Education, Entertainment and e-Management, Bali, 2011, pp. 198-201, doi: 10.1109/ICeEEM.2011.6137784.

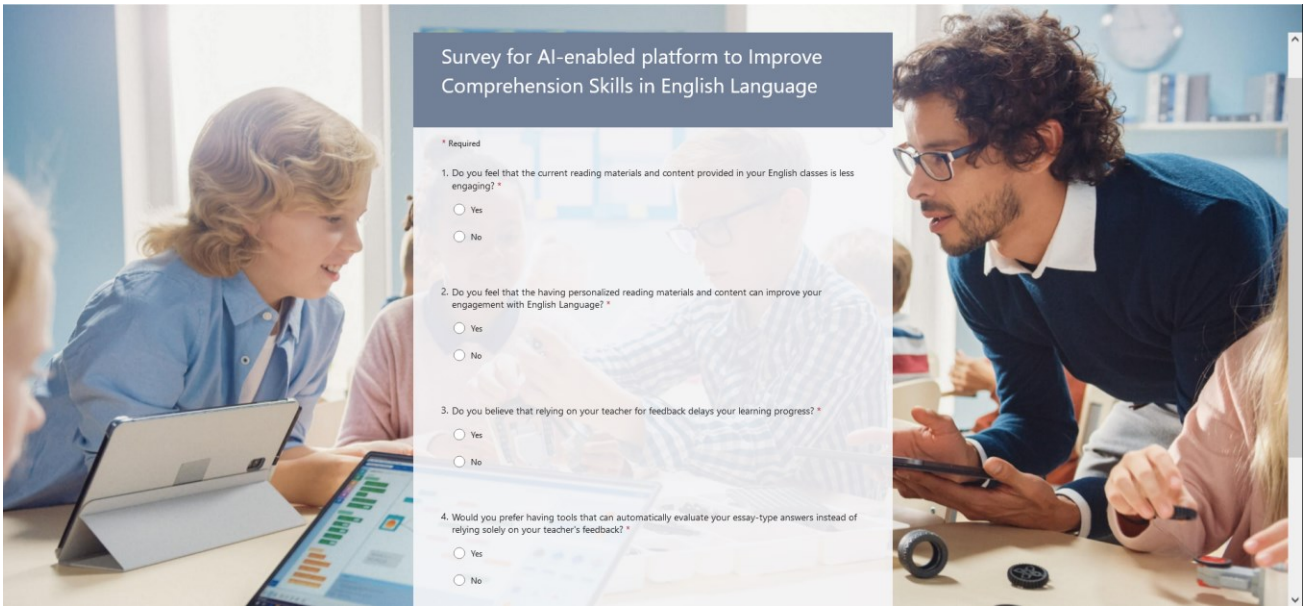
[10] R. P. Maulida, F. M. Ivone, and A. N. Wulyani, “ReadyRead: App-based Supplementary Materials for Reading Comprehension”, KSS, vol. 5, no. 3, pp. 350–364, Mar. 2021. doi: 10.18502/kss.v5i3.8557

APPENDICES

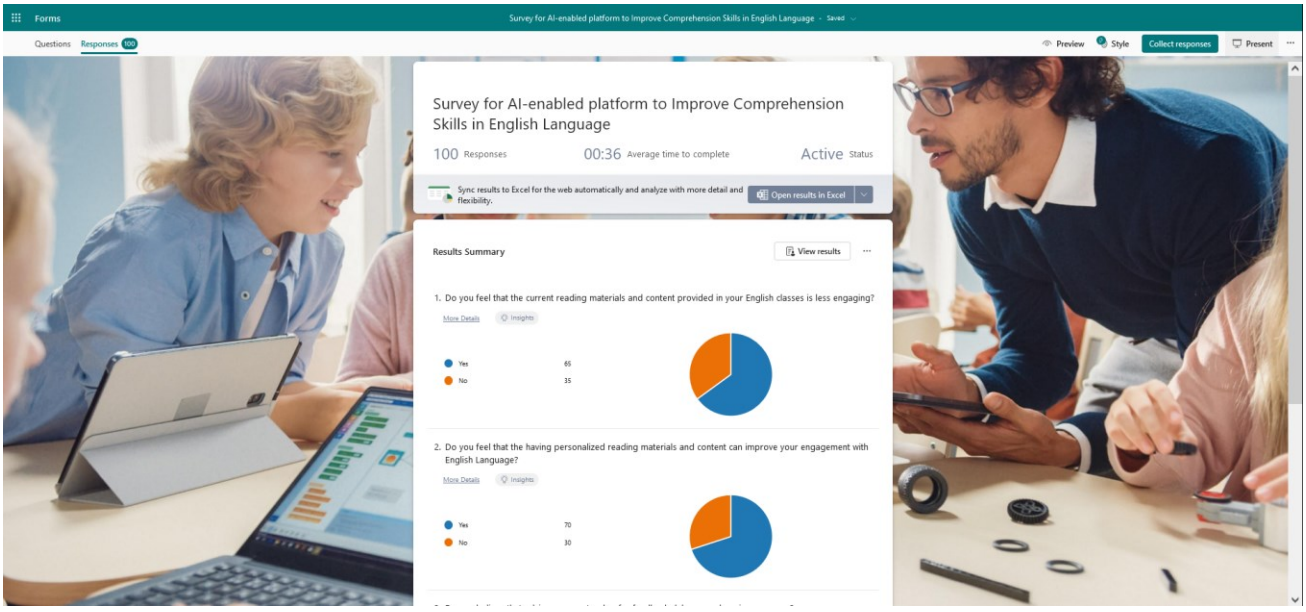
Appendix - A : Survey Questionnaire

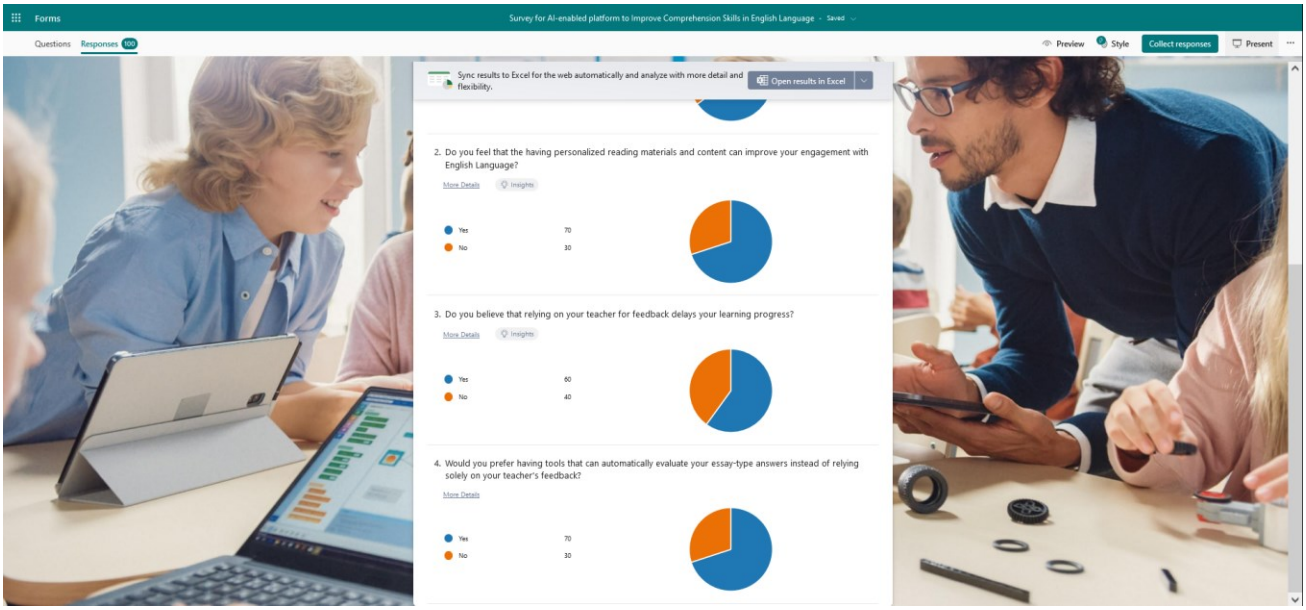
Survey Form Link - https://forms.office.com/r/UMv2qJQT8L

Survey Questions



Survey Results





Appendix - B : Plagiarism Report