SumakQuiz: An AI-powered Study Platform for Automated Table of Specifications and IRT-based Adaptive Quizzes

---

A Capstone Project Presented to
The Faculty of the College of Computer Studies
Tarlac State University
Tarlac City

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Science in Information Systems
Specialized in Business Analytics

---

by:

Clarisa T. Escoto
John Francis S. Reanzares
Mary Gaile V. Sabado
Mark L. Zablan

April 2025

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Project Context

In the rapidly evolving field of education, Artificial Intelligence (AI) has emerged as a transformative force particularly in reshaping how learning is delivered and assessed. [1], [2] Recent research has categorized AI applications in education across multiple layers such as adaptive learning, classification, recommendation, and affective computing highlighting the field's rapid advancement and growing potential for personalized instruction. [1] Traditional assessment methods which are often static, uniform, and inflexible fail to accommodate the diverse learning paces, cognitive abilities, and engagement needs of individual students. This has led to a growing demand for intelligent educational tools that support personalized and adaptive learning experiences. [1], [2], [3]

AI-enhanced quiz platforms, especially those powered by Natural Language Processing (NLP) and large language models like ChatGPT offer a promising solution.[4] Although these systems are still evolving in academic capability [5], [6], can dynamically generate questions aligned with various cognitive levels such as recall, understanding, and application while providing real-time feedback to learners. [7], [8] By leveraging frameworks like Bloom's Taxonomy and Table of Specification (ToS) such platforms ensure that generated assessments are pedagogically sound and appropriately challenging.

Despite these advances, many existing quiz systems remain limited in scope. They rely on static question banks, lack adaptive difficulty adjustment, and do not provide

personalized feedback tailored to individual learning profiles. [9] Students are often left without tools that guide their learning through formative assessment or help them track their mastery of specific subtopics. [10]

This project addresses these limitations by developing SumakQuiz, an AI-powered student-centric quiz platform designed to promote personalized formative assessment. Unlike conventional tools that depend on static question banks or instructor-authored content, SumakQuiz allows learners to upload their own study materials and generate a structured Table of Specification (ToS) limited to Lower-Order Thinking Skills (LOTS), focusing on remembering and understanding. [4], [10] Based on this ToS, the platform generates per-subtopic quizzes composed of multiple-choice questions categorized by topic and cognitive level.

Rather than adapting quiz difficulty in real-time, SumakQuiz employs staged post-quiz adaptivity. After students complete all subtopic-level quizzes, the platform uses simplified Item Response Theory (1PL model) to estimate their ability level (θ) per subtopic. [7] These ability estimates are then used to generate follow-up adaptive quizzes, adjusting question difficulty within the LOTS framework to match each learner's demonstrated proficiency. [10] This two-stage approach ensures that students are assessed at an appropriate challenge level, reinforcing comprehension while maintaining engagement.

By combining AI-driven quiz generation with adaptive questioning and personalized feedback, SumakQuiz empowers students to take control of their own

learning process. The platform not only enhances engagement and motivation but also supports deeper understanding and retention. In doing so, it represents a meaningful step toward the future of fully personalized digital education.

1.2 Purpose and Description

The purpose of this project is to develop SumakQuiz, a web-based AI-powered platform that functions as both a personalized learning assistant and a formative assessment tool for individual students. Unlike traditional quiz systems that rely on static question banks or instructor input, SumakQuiz empowers learners to self-direct their assessments using their own study materials. The platform is designed to increase engagement, foster independent learning, and reinforce foundational understanding through dynamic, AI-generated quizzes and personalized feedback.

Students begin by uploading lecture materials in PDF format and an Outcome-Based Teaching and Learning (OBTL) document. If an OBTL file is not provided, students can manually input the number of instructional hours per topic, which the system uses to support the automatic creation of a Table of Specification (ToS). This ToS is limited to Lower-Order Thinking Skills (LOTS), focusing specifically on remembering and understanding, in line with the platform's emphasis on basic comprehension and concept reinforcement.

From this structured ToS, SumakQuiz generates an initial quiz containing up to 20 multiple-choice questions, each assigned a 1-minute timer. Questions are categorized and

delivered on a per-subtopic basis, and students must complete all subtopic quizzes before proceeding to any adaptive follow-up assessments.

Once all initial quizzes are completed, the system analyzes student performance using simplified principles of Item Response Theory (1PL model) to estimate ability levels for each subtopic. Based on these estimates, the platform generates a follow-up adaptive quiz that modifies question difficulty within the LOTS framework to match the student's demonstrated proficiency. This post-quiz adaptivity ensures that learners are consistently working within their optimal learning zone.

To support repetition and retention, SumakQuiz offers a quiz regeneration feature, which allows students to regenerate each quiz up to three times. These regenerated versions do not introduce new questions but instead reword or rephrase existing items to promote active recall and varied exposure to the same concepts without altering quiz structure or difficulty mapping.

The platform includes a student-focused dashboard that visualizes performance metrics such as topic mastery, estimated ability scores, and response time. This feedback loop promotes metacognitive awareness, encouraging learners to monitor their progress, identify areas for review, and reflect on their learning path.

As a formative learning tool, SumakQuiz does not assign grades or conduct summative evaluations. Instead, it supports self-regulated learning by offering students structured opportunities for low-stakes practice, personalized feedback, and adaptive learning experiences. By combining AI-powered content processing, structured quiz logic,

and individualized assessment strategies, SumakQuiz bridges the gap between passive content consumption and active, student-driven learning.

1.3 Objectives

The general objective is to develop a student-centered AI-powered learning assistant quiz platform that dynamically generates personalized and adaptive quizzes based on user-uploaded learning materials that enable self-assessment, real-time feedback, and mastery tracking without the need for teacher involvement.

Specifically, the project aims to:

1. **To design and develop** a web-based platform that provides the following functionalities:
    1.1. User registration and authentication
    1.2. Course management interface
    1.3. Uploading of lecture materials and Outcome-Based Teaching and Learning (OBTL) documents
    1.4. AI-powered content analyzer powered by the ChatGPT API
    1.5. Automatic generation of a Table of Specification (ToS) focused on Lower-Order Thinking Skills (LOTS)
    1.6. Quiz generation system limited to 20 multiple-choice questions per upload, with a 1-minute timer per item

1.7. Subtopic-level adaptive quiz engine using simplified Item Response Theory (1PL model) for post-quiz ability estimation

1.8. Quiz regeneration functionality (up to three times) with reworded versions of the same questions

1.9. Dynamic item bank

1.10. Student dashboard

1.11. AI-powered feedback for generating personalized, formative insights based on quiz results

2. **To evaluate** the performance of SumakQuiz in terms of:

    2.1. Functionality Testing

    2.2. Accuracy Testing

    2.3. User Experience Evaluation

1.4 Scope and Limitations

This project focuses on the design and development of SumakQuiz, a web-based AI-powered quiz platform created exclusively for individual student use. The platform enables students to upload Outcome-Based Teaching and Learning (OBTL) documents and lecture materials in PDF format. If an OBTL file is not available, students may manually enter the number of hours per topic, which the system uses to support the automatic generation of a Table of Specification (ToS). The ToS is strictly limited to Lower-Order

Thinking Skills (LOTS) specifically remembering and understanding in alignment with the platform's formative and foundational learning objectives.

Although Bloom’s Taxonomy traditionally includes six cognitive levels which are remembering, understanding, applying, analyzing, evaluating, and creating, SumakQuiz currently focuses exclusively on LOTS. As a result, questions are not stratified into multiple cognitive tiers. Instead, question difficulty is handled adaptively using simplified principles of the Item Response Theory (IRT). The platform estimates each student’s ability level based on performance and adjusts the challenge level of LOTS-based questions accordingly. This allows the system to personalize difficulty based on student proficiency, rather than relying on Bloom’s cognitive classifications.

Using the ToS, SumakQuiz leverages the ChatGPT API to generate multiple-choice questions categorized per topic. Each quiz is configured with a maximum of 20 questions per upload, and each item is constrained by a 1-minute time limit. These questions are stored in a dynamic item bank and are initially delivered in per-subtopic quizzes, promoting focused, topic-specific assessment.

After completing all subtopic-level quizzes, the system performs adaptive quiz generation per subtopic, analyzing performance data through the 1PL IRT model to adjust the difficulty of subsequent quizzes to the student’s demonstrated ability level. This staged adaptivity ensures that learners are continuously challenged at an appropriate level that enhance engagement and support retention.

To encourage active recall and repetition, quiz regeneration is allowed up to three times per set, but only after all subtopic-level quizzes are completed. Regenerated quizzes do not introduce new questions but instead rephrase or reword existing ones which offers additional exposure without compromising consistency or structure.

The platform includes a student dashboard that displays topic-level mastery, estimated ability scores, and progress. As a formative assessment tool, SumakQuiz provides personalized feedback that supports self-regulated learning and metacognitive reflection. The system is intentionally built without instructor accounts, manual content validation, or administrative roles, and assumes that uploaded materials are accurate and aligned with curricular objectives.

Despite its capabilities, SumakQuiz has defined limitations. It is designed as a single-user platform, without support for collaboration, peer interaction, or educator involvement. Quiz generation is currently limited to LOTS and does not support higher-order thinking skills such as analysis, evaluation, or creation. The platform only accepts English-language, text-based PDF files and does not process scanned images, handwritten documents, or materials in other languages. As a fully web-based application, it requires a stable internet connection. Finally, SumakQuiz is intended solely for formative learning and does not offer summative grading or official academic evaluations.

These clearly defined boundaries ensure that SumakQuiz remains focused, accessible, and practical optimized to serve as a personalized assistant for quiz creation that promote foundational understanding, active learning, and student autonomy.

2 REVIEW OF RELATED LITERATURE

2.1 Discussion of Model

This section outlines the educational frameworks and computational models that inform the design and functionality of the SumakQuiz platform. These models collectively support the platform's goal of delivering personalized, formative, and self-directed assessments through AI-powered quiz generation, ability estimation, and feedback. By integrating outcome-based instructional theory with lightweight, data-informed adaptivity, SumakQuiz offers a meaningful and practical learner-centered assessment experience.

2.1.1 Outcome-Based Teaching and Learning (OBTL)

OBTL serves as the primary pedagogical foundation of SumakQuiz. The platform emphasizes aligning quizzes with clearly defined learning outcomes and performance indicators. Students are encouraged to upload OBTL-aligned documents, or alternatively, input topic-level instructional hours manually. This flexibility ensures that the platform can generate a structured Table of Specification (ToS) that maps topics and subtopics to learning goals, enhancing the educational relevance of the quizzes and supporting the principles of outcome-based curriculum design.

2.1.2 Revised Bloom's Taxonomy

SumakQuiz uses Revised Bloom's Taxonomy as a pedagogical guide for classifying the cognitive level of its quiz questions. Bloom's framework categorizes cognitive processes into six hierarchical levels: remembering, understanding, applying,

analyzing, evaluating, and creating. These are commonly grouped into Lower-Order Thinking Skills (LOTS) which are remembering and understanding, and Higher-Order Thinking Skills (HOTS) which are applying and above.

In its current implementation, SumakQuiz is intentionally limited to generating questions that target LOTS only, with a focus on foundational comprehension and factual recall. This decision aligns with the platform's role as a formative, student-directed learning assistant and reflects the current stage of system development. As a result, the platform does not stratify cognitive difficulty across multiple Bloom's levels or tiers. All questions fall within the first two levels of the taxonomy and are designed to reinforce basic understanding of uploaded learning materials.

Although Bloom's Taxonomy is not used to assign difficulty levels in the current version of the platform, difficulty is introduced adaptively through Item Response Theory (IRT). The platform estimates a student's ability level based on quiz performance and delivers follow-up questions that are statistically easier or harder while remaining within the LOTS cognitive domain. This approach ensures that students are continually challenged at an appropriate level, not by increasing cognitive complexity, but by adjusting item difficulty relative to their demonstrated proficiency.

The system's Table of Specification (ToS) leverages Bloom's classification to ensure that questions are pedagogically grounded, appropriately distributed across subtopics, and aligned with instructional objectives. This integration ensures consistency

with outcome-based educational design, even while cognitive scope is intentionally constrained to LOTS.

2.1.3 ChatGPT API as Generative Core

The ChatGPT API is the core AI engine responsible for content understanding and quiz generation. When a student uploads lecture materials, the platform uses natural language processing (NLP) to extract topics, subtopics, and key concepts. From this analysis, the API generates multiple-choice questions aligned with the content and structure defined in the ToS. Each quiz is limited to 20 questions per upload, and each question is assigned a 1-minute time limit to encourage focused, time-bound engagement. The ChatGPT API ensures semantic coherence and variation in phrasing, helping reinforce understanding through varied expression while maintaining content fidelity.

2.1.4 Adaptive Assessment via Ability Estimation

SumakQuiz incorporates adaptive assessment to personalize learning and ensure appropriate challenge for each student. This adaptivity is applied after the initial round of subtopic quizzes, not in real-time during the quiz. Once students complete all initial quizzes, the system estimates their ability level (θ) for each subtopic using principles from the Item Response Theory (IRT) specifically the One-Parameter Logistic (1PL) model, also known as the Rasch model.

The 1PL model assumes that the probability of a correct response depends on two factors of student's ability level (θ), and the difficulty of the item (b).

Unlike more complex models (2PL or 3PL), the 1PL model does not consider item discrimination or guessing, making it well-suited for formative, single-user applications like SumakQuiz. By estimating ability based on a student's responses to LOTS-level multiple-choice questions, the system can select follow-up questions that are slightly more or less difficult thus tailoring future quizzes to the learner's proficiency within each subtopic. This post-quiz adaptivity ensures that learners remain challenged without being overwhelmed and reinforces their engagement with content that matches their current level of understanding.

2.1.5 Quiz Regeneration and Active Recall

To support active recall and reinforce learning, SumakQuiz allows for quiz regeneration. After completing all initial subtopic quizzes, students may regenerate each quiz up to three times. Regenerated quizzes do not introduce new content but instead present rephrased versions of the same questions. This promotes repeated retrieval of the same concepts in different formats, supporting spaced repetition and deeper encoding of information without compromising consistency or fairness.

2.1.6 Performance Feedback and Analytics Dashboard

The platform includes a student-focused analytics dashboard that visualizes key learning metrics such as topic-level mastery, quiz history, and estimated ability scores. Metrics include per-subtopic correctness rates, average response time, and estimated proficiency levels (θ). This performance feedback loop fosters metacognitive reflection,

enabling students to track their progress, identify weak areas, and adjust their study strategies accordingly without requiring intervention from instructors.

2.1.7 Dynamic Item Bank and Topic Structuring

SumakQuiz features a dynamic item bank that plays a central role in supporting its adaptive assessment and quiz regeneration capabilities. Each multiple-choice question generated by the system is stored along with metadata such as the associated subtopic, learning outcome, cognitive level (based on Bloom's Taxonomy), and difficulty parameter estimated using the Item Response Theory (IRT) model. This structured item bank ensures that quizzes remain pedagogically aligned while offering flexibility in question selection, rephrasing, and future use. Whenever a learner uploads lecture materials, the system utilizes the ChatGPT API to extract key concepts and organize them into a hierarchical structure of topics and subtopics. The questions generated from this content are then indexed within the item bank based on their instructional alignment and position within the platform's Table of Specification (ToS).

The dynamic item bank supports several core functionalities of SumakQuiz. It enables quiz regeneration by retrieving and rephrasing previously generated questions, promoting active recall and spaced repetition without introducing inconsistencies. It also facilitates adaptive item selection by allowing the platform to draw from a pool of items that match a learner's estimated ability level, ensuring that follow-up quizzes are appropriately challenging.

2.1.8 Simplified Item Evaluation

Although SumakQuiz does not perform classical item analysis such as calculating p-values, discrimination indices, distractor efficiency, it still tracks item-level performance indicators such as per-item correctness rate, average response time, and topic-level accuracy trends. These indicators are used to refine the platform's adaptive logic and feedback, serving the individual learner's formative experience. Since SumakQuiz is designed as a single-user, formative tool, it does not require the psychometric validation processes used in summative, population-level assessments.

Table 1. Functionality and Feature Matrix

Software	Khan Academy	Quizizz	Quizlet	Gizmo	Anki
Upload PDF File content	✗	✗	✗	✗	✗
Automatic quiz generation	✗	✓	✗	✗	✗
ToS (Table of Specification) generation	✗	✗	✗	✗	✗
Alignment with OBTL	✗	✗	✗	✗	✗
AI-powered question creation	✗	✗	✗	✗	✗
IRT-based Adaptive difficulty	✗	✗	✗	✗	✓
Active recall/spaced repetition	✓	✗	✓	✓	✓
Student performance dashboard	✓	✓	✗	✗	✗
Topic-based tracking	✓	✗	✗	✗	✗
Question bank/item bank	✗	✓	✓	✗	✓

Socrative	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓
EdApp	✗	✓	✗	✗	✗	✓	✗	✓	✓	✓
SumakQuiz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

The table presents a comparative analysis between the proposed platform and existing online learning and quiz platforms such as Khan Academy, Quizizz, Quizlet, Gizmo, Anki, Socrative, and EdApp. Unlike these conventional tools, the proposed platform uniquely integrates AI-powered features including automatic multiple-choice quiz generation, dynamic item banking, post-assessment adaptive quiz creation, and Table of Specification (ToS) generation derived from user-uploaded lecture materials and optionally from Outcome-Based Teaching and Learning (OBTL) documents or manually entered instructional hours. The platform is intentionally designed for individualized, student-driven use without instructor input, offering subtopic-specific assessments, performance analytics, and personalized feedback based on simplified Item Response Theory (1PL) principles. This alignment with outcome-based instructional frameworks and its support for real-time, formative self-assessment distinguishes the platform from others,

which typically lack instructional model integration, subtopic-level adaptation, or regeneration features for repeated practice.

Table 2. Review of Related Literature

REFERENCE	DESCRIPTION	STRENGTH	WEAKNESS
S. Gupta, N. Ojeh, B. Sa, M. A. A. Majumder, K. Singh, and O. P. Adams, "Use of an adaptive e-learning platform as a formative assessment tool in the cardiovascular system course component of an MBBS programme" [11]	This study investigated the effectiveness of Firecracker, an adaptive e-learning platform, as a formative assessment tool within the Cardiovascular System course of an MBBS programme. Ninety-one second-year medical students participated in the study, completing weekly formative quizzes delivered through the platform. The research examined quiz usage rates, performance outcomes, correlations with summative exam results, and student satisfaction. The results showed that students who engaged with more quiz sessions performed significantly better in both formative and summative assessments. The study concluded that Firecracker was well-	The study provides strong evidence that adaptive, quiz-based formative tools enhance academic performance and learner satisfaction. It supports the effectiveness of personalized, data-driven assessments in reinforcing content mastery and promoting self-regulated learning.	The findings are limited to a medical education context, reducing generalizability to other fields or student groups. Additionally, the use of a proprietary platform with undisclosed adaptivity mechanisms limits insight into how similar results could be achieved using open or custom-built systems.

	received, complemented formal coursework, and contributed to improved academic outcomes.		
I. Gligorea, M. Cioca, R. Oancea, A.-T. Gorski, H. Gorski, and P. Tudorache, "Adaptive learning using artificial intelligence in e-learning: A literature review" [12]	This literature review explores the integration of artificial intelligence (AI) and machine learning (ML) in adaptive e-learning systems. Analyzing 63 peer-reviewed studies published since 2010, the review documents how AI/ML technologies—such as reinforcement learning, deep learning, and clustering algorithms—are applied to personalize learning experiences, optimize learning paths, and improve engagement, retention, and academic performance. The findings emphasize the transformative potential of AI-powered adaptive systems in shifting away from one-size-fits-all instruction toward individualized learning pathways with	This study provides a comprehensive synthesis of current research on adaptive learning technologies, clearly demonstrating how AI/ML enables content personalization and learner-specific pathways. It reinforces the value of adaptive quiz systems and performance-driven feedback by showing improved engagement and learning outcomes across multiple implementations. Its inclusion of specific algorithms and best practices, such as real-time adjustment and gamification, offers a strong theoretical and practical basis for implementing intelligent, student-	

	real-time feedback and content adjustment.	centered e-learning systems.	
V. M. IONESCU and M. C. ENESCU, "Using ChatGPT for Generating and Evaluating Online Tests" [13]	This study explores the application of ChatGPT-3 for automating both the generation of quiz questions and the evaluation of essay-type answers. It addresses the time-consuming nature of manually creating online tests and highlights the limitations of traditional test formats in terms of efficiency and vulnerability to academic dishonesty. The researchers implemented a prototype system using NodeJS, demonstrating how AI can streamline quiz content development and assessment. The paper discusses initial outcomes and outlines the technical and ethical challenges encountered in using large language models for educational testing.	The study directly supports the feasibility of using ChatGPT as a tool for quiz content generation, aligning well with systems that aim to automate assessments. It contributes to the growing body of evidence that natural language models can be leveraged to produce relevant, contextualized questions with minimal human input. The focus on implementation, outcomes, and challenges provides practical insight into integrating ChatGPT into educational platforms and highlights its potential to reduce educator workload and enhance scalability in digital assessments.	While the study showcases the utility of ChatGPT in generating quizzes, it emphasizes essay-type evaluation and does not provide detailed analysis of multiple-choice question quality or cognitive alignment with LOTS/HOTS. Further research is needed to evaluate the pedagogical reliability of AI-generated items and their role in formative versus summative assessments.

J. Firth, I. Rivers, And J. Boyle, “A Systematic Review Of Interleaving As A Concept Learning Strategy” [14]	The systematic review on interleaving as a learning strategy explores how alternating the order of examples of concepts can enhance learning, memory retention, and transfer of knowledge. The review compiles findings from 26 studies with a meta-analysis of 17 studies providing quantitative support for the benefits of interleaving. It found that interleaving boosts memory and transfer, particularly in tasks where learners need to apply knowledge to new situations, such as in art and science domains. The strategy is shown to be especially useful when distinguishing between subtly different examples and is effective even after delays. This approach contrasts with the traditional "blocked" learning method where similar items are grouped together for study.	This study provides strong evidence that varied exposure to learning content through interleaving enhances memory retention and conceptual understanding. It supports the idea that rephrasing or reordering questions can improve recall and promote deeper learning across different content areas.	While the study provides strong empirical support for interleaving, its findings are largely based on traditional or controlled learning environments, suggesting the need for further research into how interleaving principles can be effectively implemented within digital or AI-driven quiz platforms.
S. Heitmann, N. Obergassel, S. Fries, A. Grund, K.	The study investigated the effectiveness of adaptive practice	The research demonstrates that adaptive practice	While the study demonstrates the benefits of

Berthold, And J. Roelle, "Adaptive Practice Quizzing In A University Lecture: A Pre-Registered Field Experiment" [15]	quizzing compared to non-adaptive quizzing and traditional note-taking. Conducted in a real university classroom setting, it involved 188 undergraduate students who participated in three weekly lectures followed by structured follow-up learning phases. The adaptive group received quiz questions tailored to their self-rated cognitive demand, aiming to maintain moderate difficulty, while the non-adaptive group answered the same set of questions in a fixed sequence. The note-taking group reviewed lecture slides and took notes. Delayed posttests were conducted one and two weeks after the sessions to assess knowledge retention, cognitive load, and performance. Results indicated that adaptive quizzing significantly improved long-term retention and reduced intrinsic cognitive load that supports its potential as a practical and scalable alternative to more	quizzing, when tailored to individual cognitive demands, leads to better long-term knowledge retention and reduced cognitive load. Its real-world implementation in a classroom setting also highlights the practical benefits of adaptivity in learning.	adaptive quizzing, it also notes that adaptive participants progressed more slowly due to spending increased time on lower-level questions. This highlights a potential challenge in balancing personalized learning with broad content coverage in time-constrained settings.

	complex adaptive learning systems.		
L. Lim, S. H. Lim, And W. Y. R. Lim, "Efficacy Of An Adaptive Learning System On Course Scores" [16]	The study evaluates the performance of AdLeS, an in-house adaptive learning system developed to enhance academic outcomes, particularly for working adult learners. The research explored whether adaptive platforms could significantly impact student scores, using quantitative analysis such as Propensity Score Matching (PSM), t-tests, and effect size measurements (Cohen's d). Though the statistical results were not significant at the 95% confidence level, the study highlighted meaningful practical gains in learning through personalized pacing and targeted support.	The study showcases how personalized learning systems can provide targeted academic support, particularly for adult learners, and highlights the value of adjusting instruction based on learner profiles. Its use of statistical techniques enhances the reliability of its findings.	
Haryanto, Y. Neng-Shu, S. Hadi, M. Ali, A. Husna, And A. Wardhana, "Item Selection On The Moodle-Based Computerized	The study investigates the use of Moodle-based Computerized Adaptive Testing (CAT), which adjusts quiz questions based on the user's ability level using Item Response Theory (IRT). This	This study confirms the effectiveness of using Item Response Theory (IRT) for dynamically adjusting quiz difficulty, showing	

Adaptive Test" [17]	adaptive testing platform offers a personalized experience by tailoring the difficulty of questions according to the user's responses and providing real-time feedback. The system uses multiple-choice questions with varying difficulty levels to assess cognitive abilities. Data analysis demonstrated that the system effectively measures and adjusts difficulty levels, ensuring accurate assessments of knowledge. The results suggest that the Moodle adaptive test functions as intended provide tailored testing experience.	that computerized adaptive testing can accurately match item difficulty to student ability. The use of Moodle also suggests a scalable and cost-effective solution for adaptive assessment.	
J. Xu, A. Wu, C. Filip, Z. Patel, S. R. Bernstein, R. Tanveer, H. Syed, And T. Kotroczo, "Active Recall Strategies Associated With Academic Achievement In Young Adults: A Systematic Review" [18]	This systematic review examines the effectiveness of active recall-based learning strategies specifically retrieval practice (practice testing) and concept mapping in improving academic performance and self-efficacy among higher education students. By synthesizing peer-reviewed studies from	The review highlights the strong connection between active recall strategies such as retrieval practice and improved academic outcomes. Techniques like reworded quizzes and repeated exposure were shown to boost	While the study identifies the effectiveness of these strategies, it primarily focuses on traditional learning environments and underexplored digital implementation, indicating a need for further research on

	various academic databases, the review emphasized how these strategies contribute to better knowledge retention and academic outcomes. The findings emphasize that while flashcards are widely used and linked to higher GPA and test scores, other effective techniques like retrieval practice and concept mapping remain underutilized despite their positive effects on student confidence and comprehension.	retention and learner confidence across multiple studies.	integrating active recall in technology-enhanced learning systems.
N. Wengrowicz, R. Lavi, H. Kohen, and D. Dori, "Modeling with real-time informative feedback: Implementing and evaluating a new massive open online course component" [19]	This study explores the implementation and evaluation of MORTIF (Modeling with Real-Time Informative Feedback), a feature integrated into a MOOC on model-based systems engineering. MORTIF allows learners to submit models, receive detailed feedback instantly, and resubmit improved versions—enabling a "learning-by-doing" approach. The study analyzed over 12,000 submissions from 295	This study reinforces the effectiveness of immediate formative feedback combined with the opportunity to revise work, which aligns with quiz regeneration and adaptive assessment principles. It also highlights how feedback mechanisms can support diverse learning styles and foster active engagement. The	

	learners, using a mixed-methods approach combining platform analytics and surveys. Results showed increasing use of the resubmission feature and high learner preference for MORTIF compared to other question types. Students cited the immediate, actionable feedback and the ability to iterate as key contributors to improved learning outcomes.	emphasis on learner perception and iterative improvement supports the pedagogical value of systems that allow repeated, low-stakes assessment cycles—similar to quiz-based tools with regeneration or adaptation features.	
S. Liu and D. Lin, "Developing and validating an analytic rating scale for a paraphrase task" [20]	This study developed and validated a five-level analytic rating scale for assessing paraphrase tasks, a common language learning activity aimed at enhancing comprehensive language use. The researchers evaluated responses from 143 examinees using generalizability theory and many-facets Rasch analysis to confirm the scale's reliability and construct validity. The study found that paraphrasing can be assessed meaningfully across dimensions like	This study highlights that paraphrasing is not only a valid cognitive task but also assessable in a structured way, which supports the inclusion of paraphrased (reworded) questions in quiz regeneration systems. It emphasizes that rewording can reinforce understanding while preserving content integrity, aligning with the goal of concept	

	restatement, consistency, and accuracy. The rating dimensions were found to be independent, reliable, and useful for differentiating learner performance, with implications for effective feedback and classroom assessment.	reinforcement without rote repetition. The use of Rasch analysis also demonstrates how nuanced learner responses to reworded content can reflect real differences in comprehension and retention.	
R. Setyorini, "The Dominance of LOTS in Summative Assessment: The Challenge of Improving HOTS in Indonesian Language Learning" [21]	This qualitative descriptive study examines the balance between LOTS (Lower-Order Thinking Skills) and HOTS (Higher-Order Thinking Skills) in teacher-developed assessment questions within Indonesian language education. The research focused on junior high school teachers in the Brebes region of Indonesia and found that 66.9% of the questions used in summative assessments targeted LOTS, such as memorization and comprehension, while only 33.1% targeted HOTS like analysis and evaluation. The study attributes this imbalance to gaps in	The study provides valuable insight into the practical dominance and continued relevance of LOTS-based assessments in real-world educational settings. It underscores the importance of foundational cognitive skills—such as recall and understanding—which justifies platforms or tools that initially focus on LOTS-level question generation. The findings also validate the feasibility and usefulness of LOTS-focused assessments as	

	teacher training and a lack of structured guidelines for writing higher-order questions, highlighting systemic issues in assessment design and educational practice.	stepping stones toward more complex thinking skills, particularly in formative settings or where teacher capacity for HOTS design is limited.	
H. Lee and N. Bosch, "Subtopic-specific heterogeneity in computer-based learning behaviors" [22]	This study investigated how self-regulated learning (SRL) behaviors differ across various subtopics within a single domain, using data from 210 college students in an introductory statistics course. Through sequential pattern mining and coherence analysis, the researchers discovered that students engaged with metacognitive strategies—such as quizzes and feedback loops—differently depending on the subtopic. They also found that the impact of these strategies on learning gains varied, particularly in subtopics that required calculation versus conceptual understanding. The	This study provides empirical support for implementing subtopic-specific adaptivity in learning platforms, reinforcing the value of analyzing learner behavior and performance at a finer-grained level rather than treating a subject as a monolithic domain. It aligns well with systems that offer quiz generation and feedback per subtopic, validating that such granularity can capture meaningful differences in learner needs. It also advances the theoretical case for integrating personalized SRL	

	findings emphasize the heterogeneous nature of learning behavior at the subtopic level, with implications for the design of context-aware, adaptive learning systems.	strategies into computer-based environments to enhance learning effectiveness.	
J. Hua and R. Nagappan, "Research on the Influencing Factors of Students' Autonomous Learning Ability in Higher Vocational Colleges and Universities in the Context of Artificial Intelligence" [23]	This study investigates the various factors that influence students' autonomous learning capabilities in higher vocational education, particularly within the context of artificial intelligence integration. It explores how AI technologies—such as intelligent learning systems, adaptive quizzes, real-time feedback, and personalized pathways—can enhance self-directed learning. Drawing on both literature and empirical research, the study identifies critical elements affecting learning autonomy, including individual learner characteristics (motivation, prior knowledge), teacher support, and the educational environment. The research concludes that	The study strongly supports the educational value of AI-powered learning tools in enhancing autonomous, self-paced learning, which is central to your project. It emphasizes the importance of adaptive quizzes, personalized learning, and real-time feedback—all core features of your platform. Additionally, it addresses the importance of learner-centered design and acknowledges the role of AI in expanding student agency, providing a strong theoretical foundation for AI-driven assessment platforms intended to operate without	

	while AI can significantly boost autonomy and engagement, it must be implemented alongside supportive pedagogical strategies to be effective.	direct instructor involvement.	
J. S. Barrot, E. M. Masangya, and J. I. G. Lira, "Outcomes-based teaching and learning assessment instrument for teachers in higher education" [24]	This study addresses the challenge faced by higher education instructors in implementing Outcomes-Based Teaching and Learning (OBTL) effectively. It focuses on the development and validation of an assessment instrument designed to measure how faithfully in-service teachers apply OBTL principles in the classroom. Drawing data from 903 university students, the study confirmed the reliability of six interconnected domains learning process, instructional strategies, assessment, learning environment, classroom management, and teacher qualities. The instrument offers structured feedback on	The study provides strong empirical support for the core structure and components of OBTL, reinforcing the importance of aligning instructional content and assessment strategies with intended learning outcomes. This foundation justifies the integration of a Table of Specification (ToS) in systems that generate assessment content based on defined performance indicators and instructional time allocation. It also highlights the growing institutional focus on OBTL in Philippine higher education,	Although the study strengthens the theoretical basis of OBTL implementation, it is teacher-centered and does not focus on tools or systems designed for student-led assessment or autonomous learning. As such, it provides limited insight into how outcome-based frameworks can be operationalized in AI-driven or technology-mediated platforms used by students independently of instructors. Future studies could expand on how OBTL principles translate into self-directed digital

	teaching practices grounded in OBTL.	supporting the relevance of systems that are outcome-aligned.	learning environments.
B. R. B. Milloria, A. M. D. Marzonº, and L. M. C. Derasin, "Investigating AI-Integrated Instruction in Improving Academic Performance of Senior High School Students in the Philippines" [25]	This quasi-experimental study investigated the impact of integrating ChatGPT 3.5 into formative assessments among Grade 11 HUMSS students at Fervent Academy in the Philippines. Students used ChatGPT for interactive Q&A sessions during Social Sciences assessments. Pre- and post-test results showed a statistically significant improvement in academic performance, with mean scores rising from 29.25 to 43.43. The study highlighted AI's ability to enhance learning through personalized feedback, improved retention, and support for varied learning styles. It also noted key considerations related to ethics, accessibility, and the need for institutional readiness in adopting AI tools for education.	This study provides strong local evidence that integrating ChatGPT in formative assessments can significantly improve student learning outcomes. It reinforces the educational value of personalized, AI-driven feedback and adaptive content delivery—key features also present in intelligent quiz systems. Its focus on academic performance, engagement, and learning retention aligns well with the core objectives of AI-supported self-directed learning. Moreover, its Philippine context enhances the relevance of AI applications in local classrooms and contributes to the growing field of	While the study effectively demonstrates academic gains from AI use, it focuses primarily on teacher-facilitated AI interaction within a classroom setting, rather than fully student-directed platforms. This creates a gap in understanding how students might independently navigate AI-powered tools for learning and assessment. Additionally, while it identifies ethical concerns such as academic integrity and accessibility, the study does not explore in depth how these issues might be mitigated in scalable or

		AI-integrated instruction in the country.	automated systems.

Local Literature/Studies

. Barrot, Masangya, and Lira [14] developed and validated an instrument to assess how higher education instructors implement OBTL. Their findings highlight the importance of aligning assessments with intended learning outcomes, supporting SumakQuiz's automated Table of Specification (ToS) feature which is based on OBTL principles and input topic hours.

Milloria, Marzon, and Derasin [15] examined the effects of ChatGPT integration in formative assessments for senior high school students. Their results showed significant academic improvement, validating the use of AI in enhancing retention and engagement. While their setup involved teacher facilitation, the study supports SumakQuiz's goal of AI-powered, student-led assessments in Philippine educational contexts.

Foreign Literature/Studies

. Gupta et al. [1] explored the effectiveness of Firecracker, an adaptive e-learning platform used as a formative assessment tool in a medical course. Their findings indicated that students who regularly engaged with the quizzes performed better in both formative and summative assessments. This reinforces the pedagogical value of adaptive quizzing for improving academic outcomes, aligning well with SumakQuiz's use of post-quiz adaptation based on student performance.

Gligorea et al. [2] conducted a comprehensive literature review on the integration of artificial intelligence and machine learning in adaptive e-learning. They highlighted how these technologies support personalized learning pathways, real-time feedback, and dynamic content adjustment—all central to SumakQuiz's intelligent, learner-centered approach.

The feasibility of using large language models for quiz generation was investigated by Ionescu and Enescu [3], who implemented a ChatGPT-based system for creating and grading online assessments. Although focused on essay-type evaluations, the study supports the potential of ChatGPT to automate educational assessments and reduce instructional workload, which complements SumakQuiz's AI-driven quiz generation.

Firth et al. [4] reviewed the interleaving strategy in concept learning, showing its effectiveness in enhancing memory and transfer of knowledge. Their findings support SumakQuiz's regeneration feature, where quiz questions are reworded to improve retention while avoiding repetition fatigue.

In a field experiment, Heitmann et al. [5] found that adaptive quizzes tailored to students' perceived difficulty levels improved long-term retention and reduced cognitive load. This supports SumakQuiz's adaptive design, which modifies quiz difficulty per subtopic based on user performance using simplified Item Response Theory (IRT).

Lim et al. [6] evaluated AdLeS, an adaptive learning platform aimed at adult learners. Although statistical significance was limited, the study emphasized meaningful

learning gains through personalized pacing—an approach echoed in SumakQuiz's self-directed quiz structure.

Haryanto et al. [7] applied IRT within Moodle-based computerized adaptive testing. Their work validates the use of IRT to tailor question difficulty to a learner's ability level, which is a core mechanism in SumakQuiz for generating follow-up quizzes that match a student's current mastery.

A systematic review by Xu et al. [8] emphasized the effectiveness of active recall strategies like retrieval practice in boosting academic performance and self-efficacy. These findings align with SumakQuiz's regeneration and feedback mechanisms that promote concept reinforcement and knowledge retention.

Wengrowicz et al. [9] implemented MORTIF, a MOOC feature offering real-time formative feedback and iterative submission. Their study supports the educational value of systems that allow re-attempts and formative feedback—concepts central to SumakQuiz's regeneration and dashboard components.

Liu and Lin [10] developed and validated a rating scale for paraphrased responses. Their findings support the assessment value of reworded content, providing a theoretical foundation for SumakQuiz's regeneration strategy, which reconstructs questions to reinforce understanding.

Setyorini [11] highlighted that 66.9% of teacher-created assessments in Indonesia focused on Lower-Order Thinking Skills (LOTS). This justifies SumakQuiz's current

design, which emphasizes LOTS-based questions for foundational comprehension in formative settings.

Lee and Bosch [12] studied self-regulated learning behaviors across subtopics and found variability in student engagement and effectiveness of strategies. Their research supports SumakQuiz's subtopic-based quiz design and post-assessment adaptation, which account for varied student performance per topic.

Hua and Nagappan [13] investigated how AI impacts autonomous learning in vocational colleges. Their work affirmed that adaptive quizzes, personalized feedback, and intelligent systems foster autonomy and engagement, reinforcing the learner-centered philosophy of SumakQuiz.

2.2 Conceptual Framework

Figure 1. Conceptual Framework of SumakQuiz

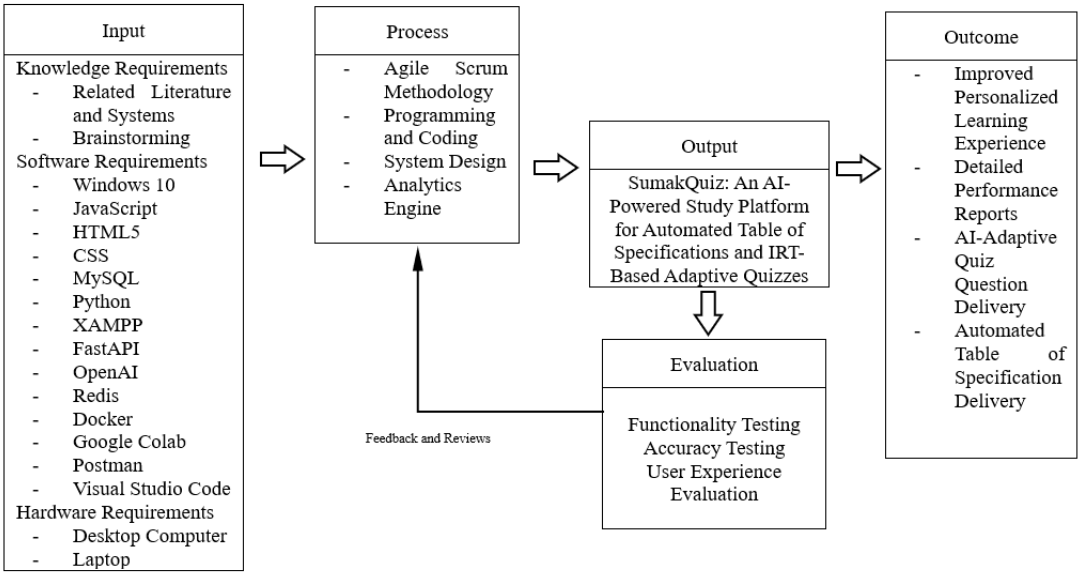


Figure 1 shows the conceptual framework which illustrates how the system, SumakQuiz, will determine its input, process, output, and outcome. This framework is based on the design established by the Input-Process-Output Model.

The input phase includes the knowledge requirements, related literatures and systems will be used to collect various insights with regards to what and how will the system functions, brainstorming sessions is also included. Windows 10, JavaScript, HTML5, CSS, MySQL, Python, XAMPP, FastAPI, OpenAI, Redis, Docker, Google Colab, Postman, and Visual Studio Code, as well as the necessary hardware such as desktop and laptop support the technical and functional foundation of SumakQuiz.

The process ensures that incremental sprints occur with continuous iteration. Additionally, it involves Agile Scrum Methodology, programming and coding, system design, and integration of an analytics engine.

The output will be the SumakQuiz study platform where testing and evaluation such as functionality, accuracy, and usability testing will be conducted ensuring the quality of the system by following the criteria outlined in ISO 25010.

Moreover, the outcome will result to delivering a system that improved personalized learning experience, support automated table of analysis generation, enable AI-adaptive quiz questions delivery, and generating detailed performance reports.

This conceptual framework serves as a graphical representation to guide the admin, users, and readers in understanding how the system will be designed and evaluated based on the project's objectives.

2.3 Definition of Terms

Adaptive Assessment

In this study, adaptive assessment refers to the system's ability to generate quizzes with varying difficulty levels based on a student's prior performance. It uses simplified principles from Item Response Theory (1PL model) to estimate ability and tailor subsequent questions per subtopic accordingly.

AI-Powered Quiz Generation

In the context of this project, this term refers to the use of the ChatGPT API to automatically generate multiple-choice questions from uploaded learning materials. The system extracts topics and concepts and formulates questions without human intervention.

ChatGPT API

Within the study, this refers to the external generative artificial intelligence engine that processes user-uploaded materials to generate relevant, LOTS-based questions and structured feedback. It acts as the core content analyzer and item generator for quizzes.

Formative Assessment

In this project, formative assessment refers to low-stakes quizzes designed to monitor student learning progress. The results are used to provide real-time feedback and guide self-directed review rather than to assign official grades.

Item Response Theory (1PL Model)

For this system, IRT refers to the computational model used to estimate a student's ability level (θ) based on quiz performance. The 1PL or Rasch model considers only the difficulty of items and correctness of responses in adapting quiz difficulty per subtopic.

Lower-Order Thinking Skills (LOTS)

In this study, LOTS encompasses the first two levels of Revised Bloom's Taxonomy—remembering and understanding. SumakQuiz is currently designed to generate and evaluate questions strictly within this cognitive domain.

Outcome-Based Teaching and Learning (OBTL)

Within the context of this project, OBTL refers to the instructional approach where assessments are aligned with defined learning outcomes. Students may upload OBTL-aligned documents or manually input instructional hours to guide the generation of the Table of Specification (ToS).

Paraphrased Quiz Regeneration

In SumakQuiz, this refers to the process of rewording or reconstructing previously generated questions for review. Students may regenerate a quiz up to three times, with all regenerated items remaining within the same conceptual framework to reinforce learning.

Self-Directed Learning

In the platform, this refers to the student's autonomy in managing their learning process—uploading materials, completing quizzes, receiving feedback, and deciding when to regenerate content without relying on an instructor.

Subtopic-Level Adaptation

This term, as used in the system, refers to the adaptive quiz generation that occurs after completing all initial quizzes. The system generates follow-up questions with adjusted difficulty specific to each subtopic based on a student's demonstrated proficiency.

Table of Specification (ToS)

In the study, the ToS is an automatically generated matrix that maps uploaded lecture content or OBTL hours to subtopics and LOTS-based question counts. It ensures quiz content is distributed proportionally and pedagogically aligned.

Quiz Regeneration

Quiz regeneration is the feature that allows students to request up to three reworded versions of a previously taken quiz. It supports active recall and spaced repetition by reinforcing the same concepts in varied linguistic formats.

Student Dashboard

In this platform, the dashboard is a visual feedback interface where students can monitor their performance trends, topic mastery levels, and estimated ability scores. It supports metacognitive reflection and personalized review.

Multiple-Choice Questions (MCQs)

In the context of SumakQuiz, MCQs are the sole format used for assessments. Each quiz includes a maximum of 20 items, with a time limit of one minute per item to standardize testing and promote focus.

Learning Analytics

Within the study, learning analytics refers to the tracking and analysis of quiz data to inform the adaptation of future content and provide personalized feedback.

3 TECHNICAL BACKGROUND

3.1 Software Development Requirements

Table 3. Software Development Requirements

Software	Description	Photo
Windows 10	An operating system which serves as the operating platform for the development of the system.	
JavaScript	A programming language to create web pages interactive.	
HTML5	A markup language used for structuring and presenting the content of the system's web pages.	
CSS	A style sheet style utilized to style and visually present the HTML structure of the web pages.	

MySQL	This software will be responsible for storing, organizing, and managing all essential data (user profiles, quiz scores, and system logs).	
Python	This will be the backend programming language which will be handling system logic, AI integration, API development, and generating the Table of Specifications (ToS) based on content analysis and OBTL inputs.	
XAMPP	XAMPP will serve as the local server environment for the development and testing of the system.	
FastAPI	A modern python web framework used for building APIs that manage ToS generation, quiz creation, and adaptive learning features.	
OpenAI	This will be used for providing AI functionalities for generating quizzes, analyzing uploaded content to categorize topics and cognitive levels, assisting in automated ToS generation.	
Redis	Storage for efficiently caching quiz responses and speeding up personalized recommendations.	

Docker	This will serve as the container for consistent development, testing, and deployment environments.	
Google Colab	This will be used during development to prototype and test the AI algorithms/data analytics.	
Postman	This will serve as a tool for testing and documenting RESTful APIs between frontend, backend, and AI services.	
Visual Studio Code	This software will serve as text editor that will be used in the development of the system.	
Pandas (Python Library)	This will be utilized for creating, managing, and exporting the Table of Specifications (ToS). This enables tabular data manipulation, counting items per topic/outcome, and exporting to Excel or CSV formats which is essential for structuring AI-generated quiz data and analytics.	

Canva	Used for creating the mockup design of the platform.	

3.2 Hardware Development Requirements

Table 4. Hardware Used for the Development of the Platform

Hardware	**Description**	**Photo**
Laptop	Huawei Matebook 14 (2021)	
Laptop Specifications		
Processor	AMD Ryzen 5 5500U with Radeon Graphics 2.10 GHz	
Memory	16 GB	
Storage	512 GB	
Operating System	Windows 11	

Table 4 shows the hardware that the researchers will be utilizing to develop the system, including the specifications of the device. The presented device is capable of coding, running, and testing the system.

Table 5. Hardware Specification Requirements

Description	PC/Laptop that will be used by the learner/instructor.
Processor	I3-1155G4 or any dual-core equivalent.

Memory	4GB
Storage	250GB
Operating System	Windows 7

Table 5 shows the minimum requirements for testing the proposed system.

Table 6. Hardware Specification Requirements

Description	Mobile Phone that will be used by the learner/instructor.
Photo	
Chipset	Tecno Camon 30 Pro
Memory	12GB RAM
Storage	256GB
Operating System	Android, IOS

Table 6 shows the specifications of the mobile device that the learner/s will be using to test the system.

3.3 Peopleware

3.3.1 Project Team

This team is composed of researchers who will be developing the proposed system. This consists of two main sections, the Documentation and Development Team. The Documentation Team oversees writing down the progress and the papers containing the research's details. While the Development Team is tasked with coding and running the system.

The researchers should have the skills and abilities required to code, run, and test the actual systems, including the documentation.

3.3.2 Stakeholders

Since this platform will primarily serve as an assistant for student learning, the main stakeholder will be the students. They will be the direct users of the system by interacting with AI-powered quiz generator that promotes learning and monitor their progress.

Additionally, institutions that implement Outcome-Based Teaching and Learning (OBTL) frameworks may benefit from the system's ability to align assessments with learning outcomes.

3.4 Sources of Data

The data for developing, testing and evaluation of the AI-powered adaptive quiz platform will mainly rely on the content provided by user, system-generated data, and academic research to provide a smart and personalized learning experience for users. The

platform will be fully student-driven, which does not require instructor interaction. Hence, students may upload their own study materials which will be the primary source for generation of quizzes. The platform will utilize natural language processing (NLP) to generate questions that are aligned with the material provided. Users must upload an OBTL framework file or they can manually input OBTL details such as the lecture title and the hours of how long the topic was taught. This information guides the AI system in generating questions aligned with the educational objectives.

Since the platform uses AI (OpenAI GPT-4) to create quizzes based on the uploaded materials and OBTL information, these generated items will be stored and analyzed to improve quality, relevance, and adaptability across different topics and difficulty levels. Also, the platform will be collecting the quiz scores, time per question, and performance per topic. This will serve as the basis for adjusting difficulty levels, and personalized learning assistant for users.

Lastly, feedback will be gathered through forms for the evaluation in terms of functionality, accuracy, and user satisfaction. The gathered results will serve as basis for improving the platform’s question generation logic, user interface, and adaptive features.

4 METHODOLOGY

4.1 Methods in Data Gathering

4.1.1 Brainstorming

The researchers aimed to leverage the potential of AI in education, addressing the inefficiencies of ordinary static systems. Existing learning platforms often lack the ability to adapt to the learner's ability level, limiting the learning outcomes and failing to retain engagement. With the rise of artificial intelligence such as ChatGPT, the researchers conceptualized using AI to power the web-based adaptive quizzing platform to help address individual student's needs in learning. The system has drawn inspiration from the principles of OBTL and dynamically generates quiz items of varying cognitive levels, particularly focusing on Lower-Order Thinking Skills (LOTS). By the utilization of generative AI and natural language processing (NLP), the platform creates personalized quizzes that adjust to a student’s performance over time. It integrates simplified Item Response Theory (IRT) for adaptive difficulty, promoting mastery learning and delivering formative feedback tailored to each learner.

4.1.2 Internet Method

The researchers utilized online academic resources to explore the development of AI-based adaptive quiz systems. Information about adaptive learning principles and personalized construction, IRT application, capabilities of ChatGPT/GPT-4 on quiz generation, and real-world examples of AI-enhanced learning platforms. Moreover, the

internet informed the researchers of the effectiveness of generative AI tools in boosting student engagement and supporting learning strategies that align with the outcome.

4.2 Story Board

In the development plan for the platform, there are several considerations and constraints on how the system is structured in terms of user experience and technical implementation.

Design Considerations:

1. The platform will tailor the question difficulty based on the student's performance using IRT and learning analytics.
2. The platform will integrate ChatGPT/GPT-4 to dynamically generate high-quality questions and explanations that align with cognitive skills.
3. The platform will dynamically adjust questions based on the student's previous performance, which still offers some challenges in answering the question without overwhelming the students.
4. The platform will utilize an automatically generated TOS based on the uploaded OBTL framework file or student-inputted number of hours taught to assist in creating questions.
5. The platform will also include an analytics dashboard, describing the student's performance and tracking their progress.
6. The platform will utilize a minimal design, prioritizing function integration over aesthetics.

Constraints:

1. The platform's AI-generated questions and feedback may not be accurate as the platform only relies on AI's understanding of the context.
2. The platform will require stable internet connectivity to access and retrieve AI-generated content and track user data.
3. The platform may require higher server resources or subscription costs to host AI services and analytics.
4. The platform must apply proper filtering and verification mechanisms to avoid bias and misleading content.
5. The platform may require fine-tuning of questions generated by ChatGPT based on local curriculums and student proficiency levels.
6. The platform's use of OpenAI's GPT models may be constrained by API quotas or cost-related throttling.

Figure 2. Mockup Design

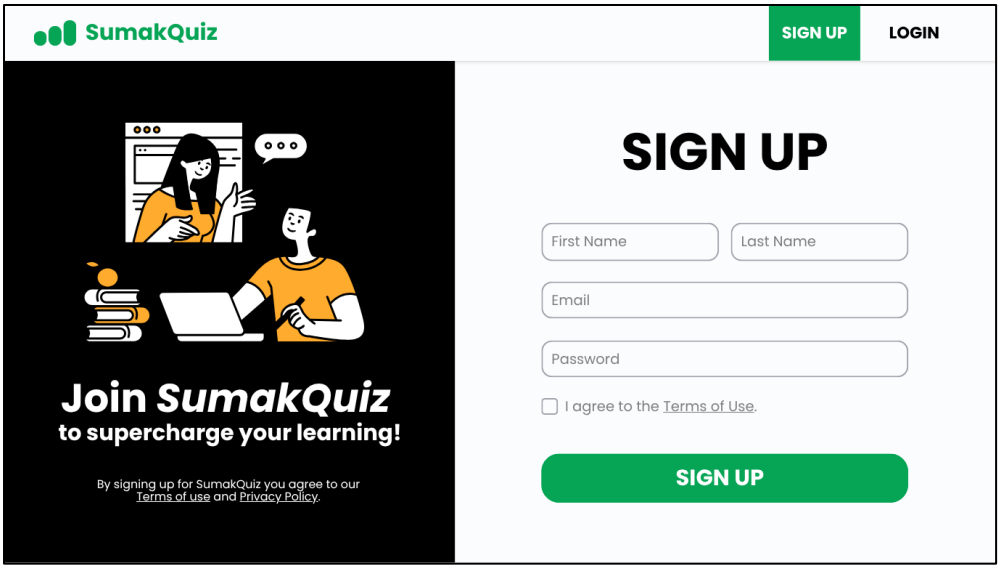
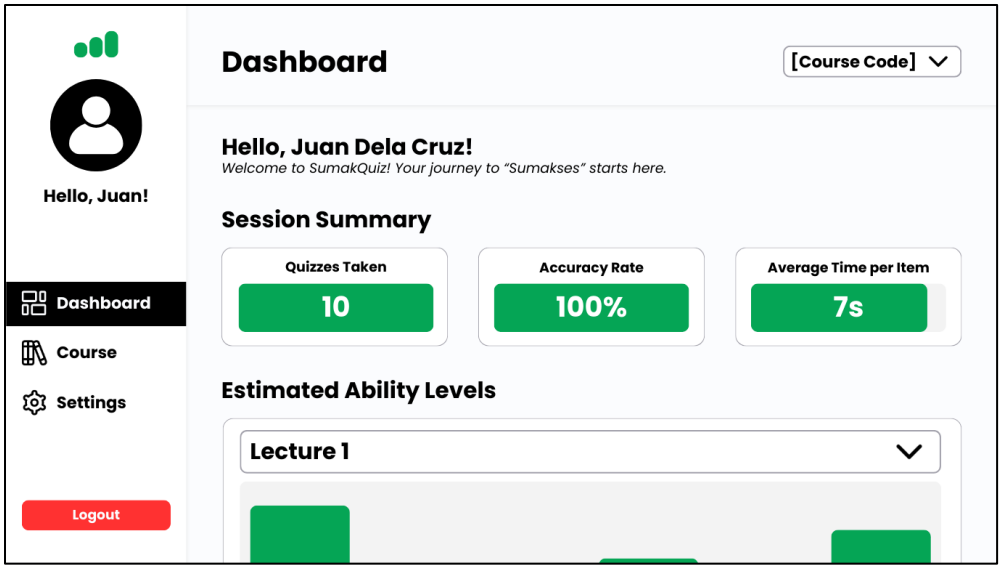
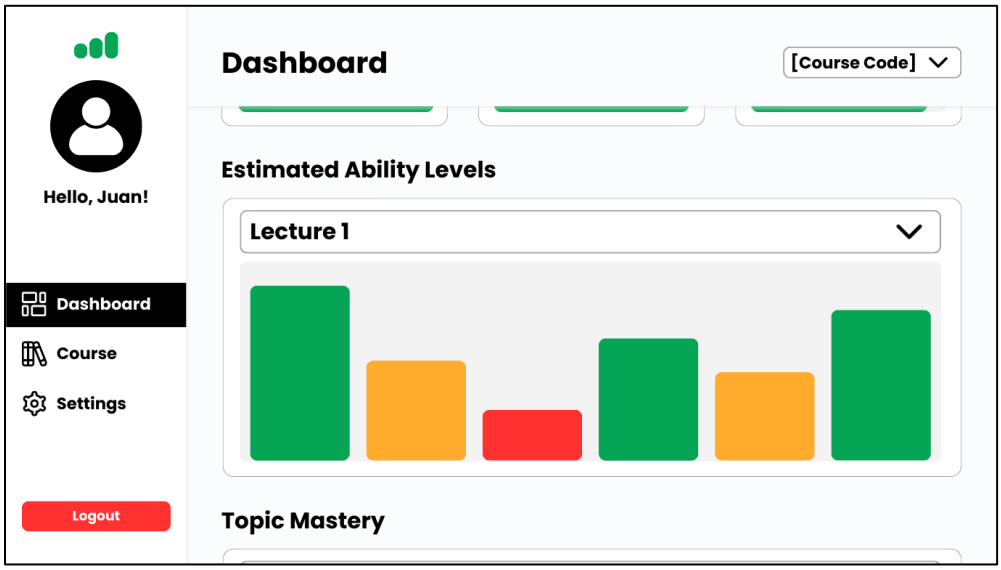


Figure 3. SumakQuiz Sign Up Page

Figure 3 shows the SumakQuiz sign up page. The sign-up page requires the following input is required: First Name, Last Name, Email, Password, and Agreement to the Terms of Use which must be read before ticking the checkbox. The Terms of Use and Privacy Policy can be read by clicking on their respecting links. After the input required is provided, the learner may sign in to the platform.


Hello, Juan!
Dashboard
Course
Settings
Logout
Dashboard
[Course Code]
Hello, Juan Dela Cruz!
Welcome to SumakQuiz! Your journey to "Sumakses" starts here.
Session Summary
Quizzes Taken
10
Accuracy Rate
100%
Average Time per Item
7s
Estimated Ability Levels
Lecture 1


Hello, Juan!
Dashboard
Course
Settings
Logout
Dashboard
[Course Code]
Estimated Ability Levels
Lecture 1
Topic Mastery

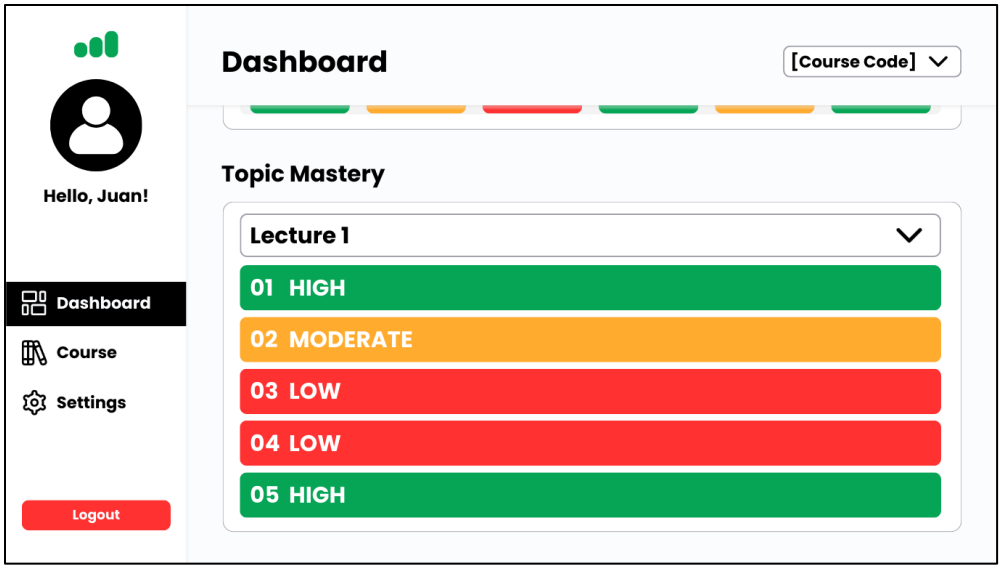


Figure 4. SumakQuiz Dashboard Page

Figure 4 shows the SumakQuiz dashboard page. This is where the user is redirected upon signing in and/or logging in. In addition, here is where the student may monitor its progress and performance. The dashboard can be viewed by course which provides a session summary presenting the total number of quizzes taken, overall accuracy rate, and average time per item; estimated ability levels which can; estimated ability levels for each subtopic; and topic mastery which shows the student's mastery of the lecture's subtopics.

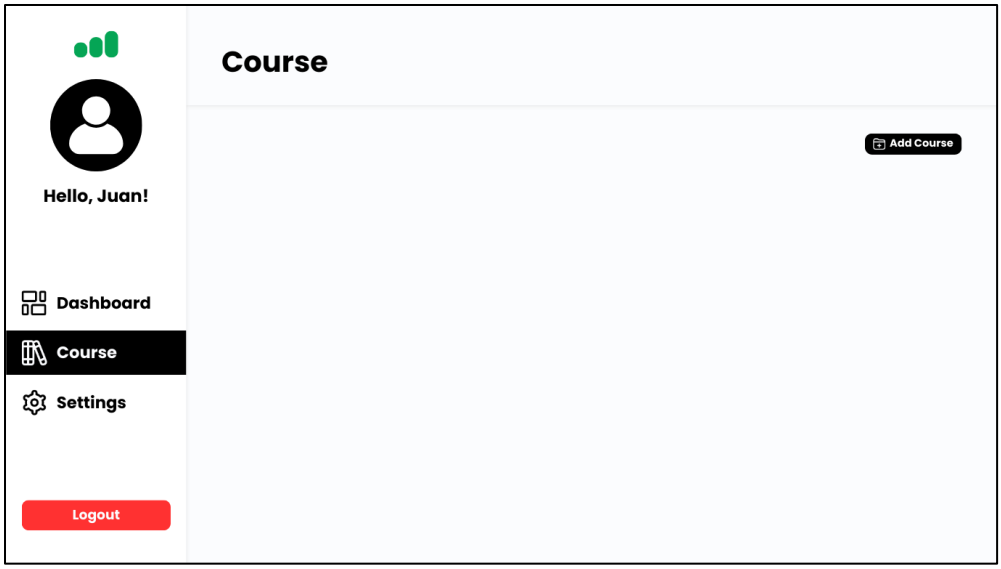


Figure 5. SumakQuiz Course Collection Page

Figure 5 shows the SumakQuiz course collection page. This page allows the user to add a course through the “Add Course” button placed at the top right of the page.

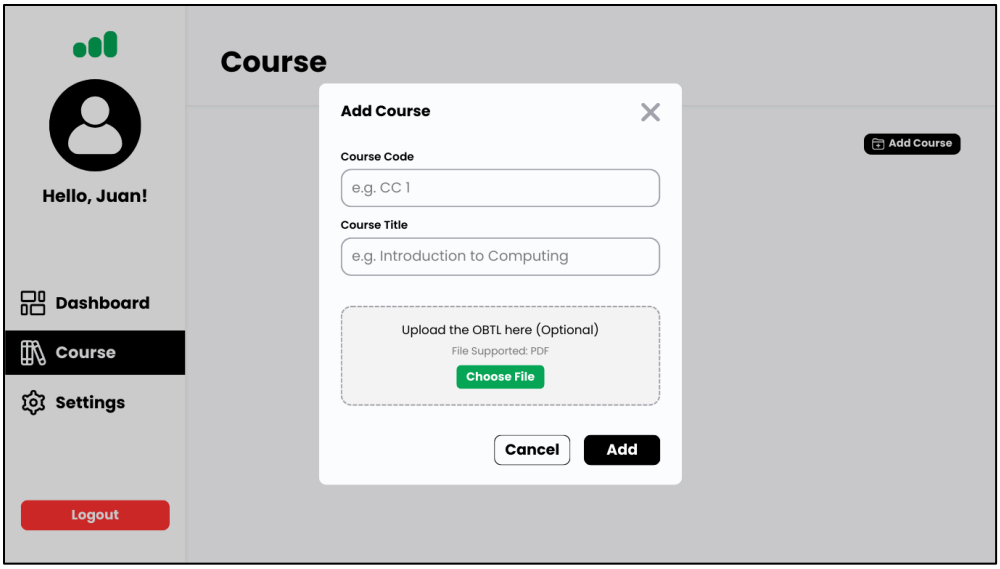


Figure 6. SumakQuiz Course Addition

Figure 6 shows the SumakQuiz course addition. This student is required to provide the following input: Course Code and Course Title. The student can optionally upload the course's OBTL. After the required inputs are encoded, the user may click on the "Add" button to add it to the Course Collection.

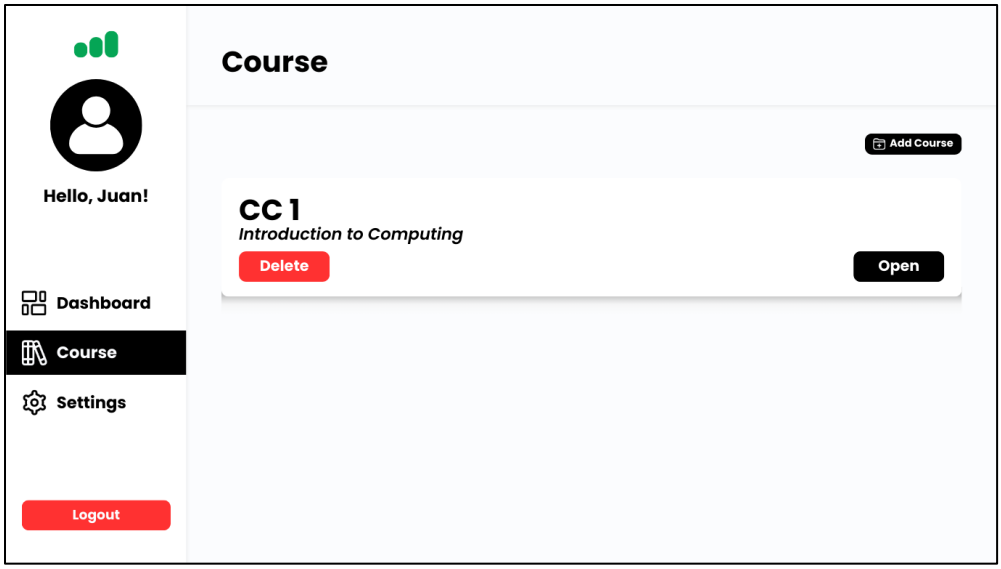


Figure 7. SumakQuiz Course Added

Figure 7 shows the SumakQuiz course added. This is what the Course Collection page will look like once a course is added. The user may delete the course and open it to generate topic quizzes.

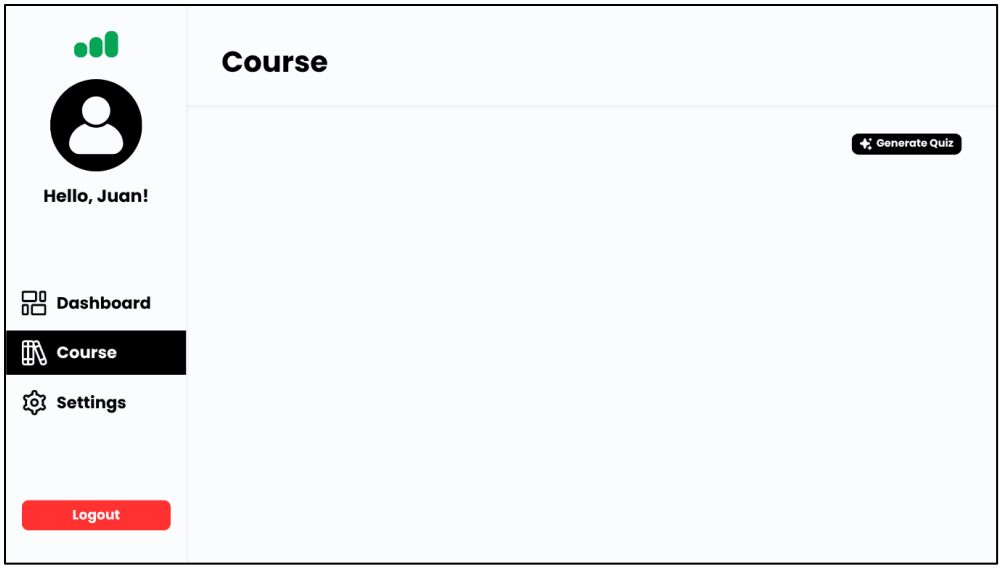


Figure 8. SumakQuiz Course Page

Figure 8 shows the SumakQuiz course page. This is what it looks like once the user opened the course without the generated quizzes. This page is where the student may generate a quiz for each topic by clicking on the “Generate Quiz” button.

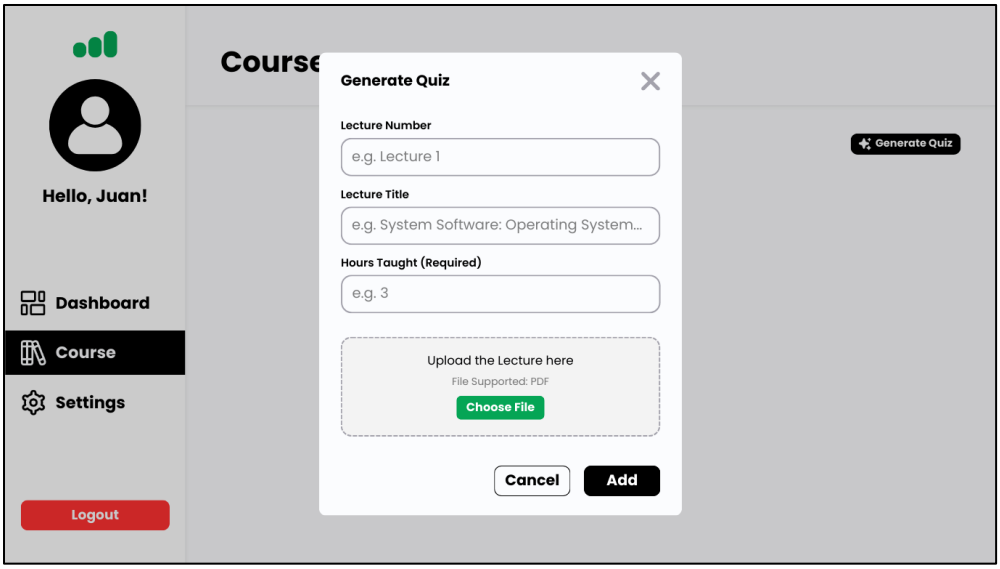


Figure 9. SumakQuiz Quiz Generation

Figure 9 shows the SumakQuiz quiz generation. The user is required to enter the lecture number, lecture title, number of hours taught is only required if the student was not able to provide the OBTL of the course and upload the lecture file. After the required inputs are encoded, the user may click on the “Add” button to generate subtopic quizzes for the lecture.

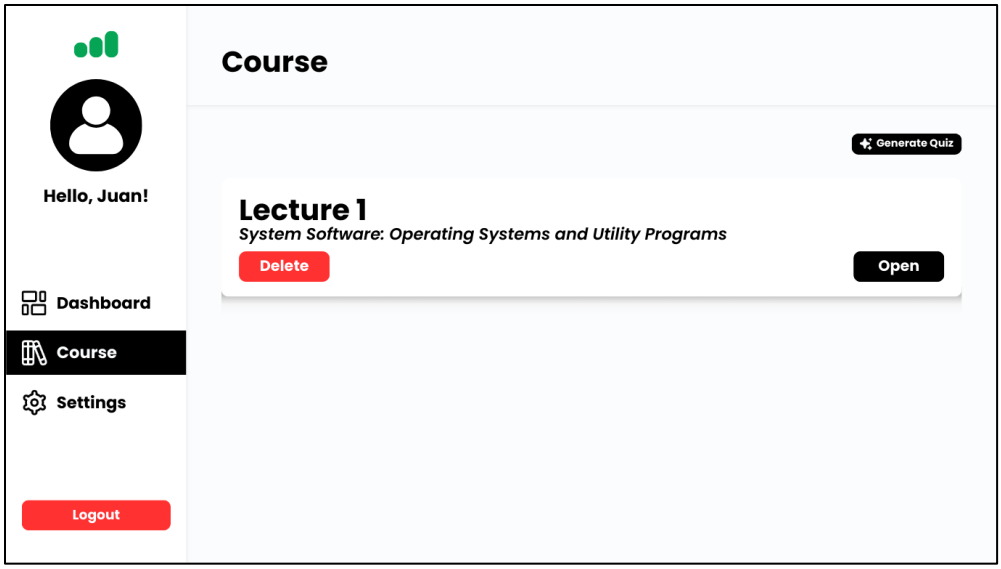


Figure 10. SumakQuiz Generated Lecture Quiz Added

Figure 10 shows the SumakQuiz generated lecture quiz added. This will be where the course lecture quiz collection will be accessible. The student may delete the lecture quiz generated. Once the student opens the lecture quizzes, they will be redirected to another page where the collection of subtopic quizzes are available for taking.

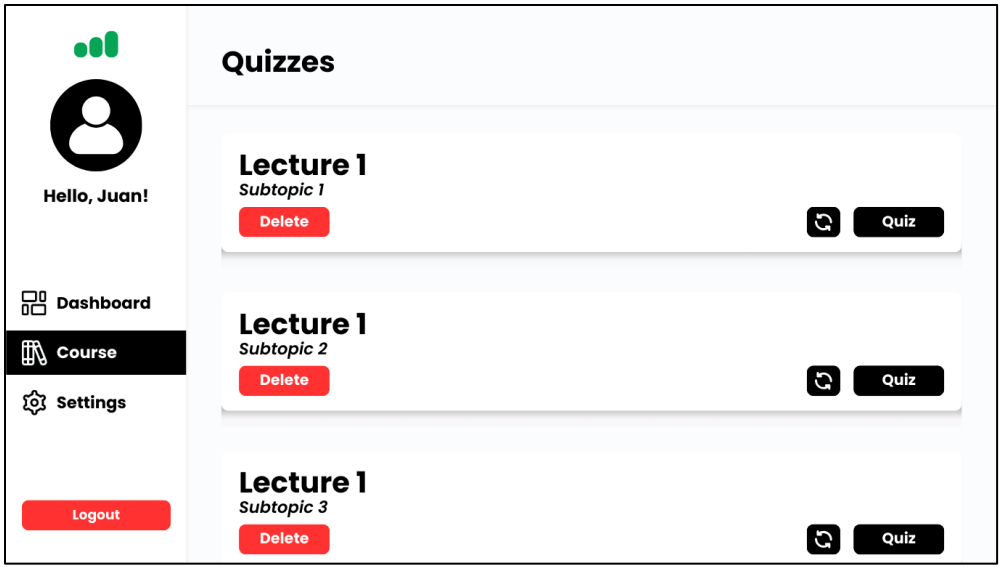


Figure 11. SumakQuiz Subtopic Quizzes Collection

Figure 11 shows the SumakQuiz subtopic quizzes collection. This is where the generated quizzes for the lecture are available for taking. The student is required to finish the first version of the quizzes before they can regenerate a maximum of two more versions of the subtopic quiz which are both adaptive. Once the student clicks on the “Quiz” button, the will directed to the start of the actual subtopic quiz.

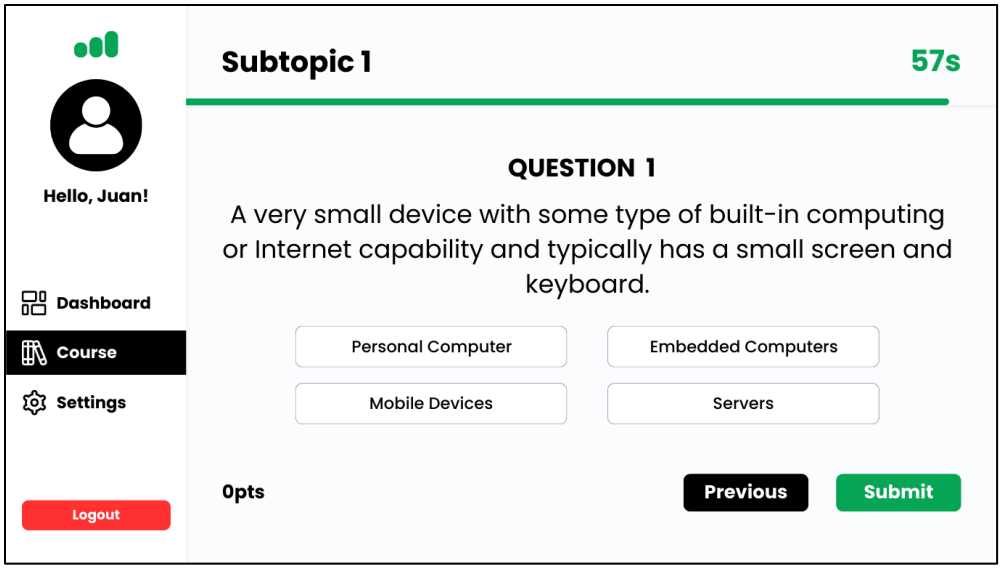


Figure 12. SumakQuiz Quiz Item Sample

Figure 12 shows the SumakQuiz quiz item sample. This is where the student is presented with a multiple-choice question. Each item must be answered in less than 60 seconds. The interface shows the count down at the upper right side of the page and colored bar that shrinks along with the count down. Both are colored green during the 60-31 seconds mark, they are colored yellow during the 30-11 seconds mark, and they are colored red during the 10-0 seconds' mark. After the student picks its answer, they must click on the "Submit" button which triggers the checking of the answer. If the selected answer is wrong the button of the selected answer will turn into red and the points count will remain the same, otherwise it will turn into green, and the points count will accumulate.

4.3 Design and Development of SumakQuiz: An AI-Powered Study Platform for Automated Table of Specifications and IRT-Based Adaptive Quizzes

4.3.1 Requirement Analysis and Documentation

4.3.1.1 User Requirements

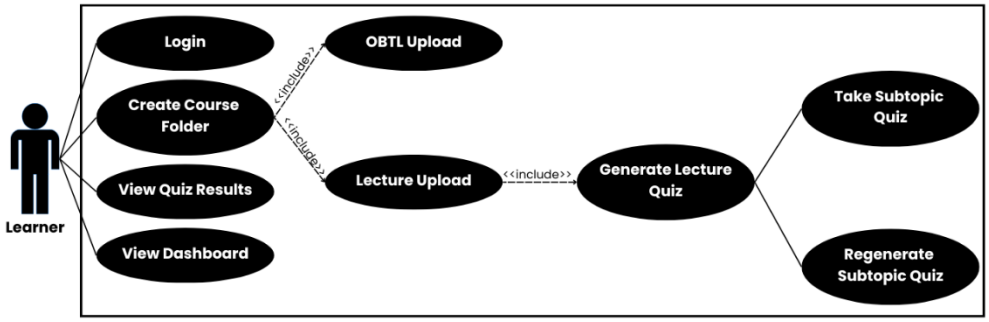


Figure 13. Use Case Diagram of SumakQuiz

Figure 13 shows the Use Case Diagram of SumakQuiz. The diagram shows the actor, the Learner. The Learner can login, create a course folder, view quiz results, and view dashboard. Moreover, upon the creation of the course folder, the learner can optionally upload the OBTL for the course. Once the course folder creation is done, the learner can upload lectures to generate quizzes. The generation of lecture quiz is divided according to the subtopics the lecture has. The learn then takes the quizzes per subtopic in a lecture. After taking the first versions of the quizzes, the learner can regenerate for a maximum of two more generations which are adaptive.

4.3.1.2 User Characteristics

The primary user of the system are the learners as the platform is mainly conceptualized to partly support them in their academic endeavor. The users are expected to be literate enough to use computers.

The learner will be able to:

1. Login
2. Create a course folder
    i. Upload the OBTL optionally
    ii. Upload the lecture
        a. Generate lecture quiz
            1. Take subtopic quiz
            2. Regenerate subtopic quiz
3. View quiz results
4. View Dashboard

4.3.1.3 Functional Requirements

The table outlines the functional requirements for the Login Module of SumakQuiz. It includes input validation, user authentication, role-based access, session management, password encryption, password recovery, and responsive user interface.

Table 7. Functional Requirement for Login Module

Req.ID	Requirement Description	Priority	Complexity
FR1	To login to the platform the system shall validate user input. If the email and password fields are filled	High	Medium

Req.ID	Requirement Description	Priority	Complexity
	out incorrectly, the system will throw an error message to the user. The same happens if there are any empty fields.		
FR2	Registered users shall be able to authenticate their identity using their login credentials. The email and password must match with a record on the database. Otherwise, the user cannot gain access to the system.	High	Medium
FR3	The system shall redirect registered users to the appropriate landing page according to the role they declared during the sign-up process.	High	Medium
FR4	After a user's successful login, a user is detected active, and a session is counted. If the user is inactive for a set of periods, the session shall expire which will automatically log the user out and require them to login again.	High	Medium
FR5	Passwords shall be protected using password hashing to make sure that the password is not stored in plaintext.	High	Medium
FR6	Users shall be allowed to reset forgotten password via email verification to ensure they can regain access securely.	Medium	Medium
FR7	The platform will offer a responsive layout which will ensure that the platform will be usable both on desktop and mobile devices.	Medium	Low

4.3.1.4 Non-Functional Requirements

The table outlines the non-functional requirements for the Safety and Security of SumakQuiz: A Study Platform with Learning Strategies and Item Response Theory Adaptive Quizzing. It includes input validation, user authentication, role-based access control, session management, password encryption, and password recovery. To uphold the platform's ethical standards in securing the user's data, the system should comply with the national data protection law, the Philippine Data Privacy Act of 2012.

Table 8. Non-Functional Requirement for Safety and Security of SumakQuiz

Code	Dependencies Description
SS1	All inputs must be validated on both the client and the server side. Any invalid input must trigger an error message along with an instruction (if necessary) without revealing any information from the system. This requirement will depend on form validation on HTML5.
SS2	The verification of the user's identity will prevent unauthorized logins ensuring only real users can login with the right credentials. This requirement will depend on the use of the built-in login/register system of Laravel Breeze.
SS3	Role-based access control restricts features that will prevent a user from accessing inappropriate information. This requirement will depend on if-checks associating with the role to route the user to an appropriate landing page.
SS4	User session management prevents session hijacking. This requirement will depend on the default session handling Laravel.
SS5	Passwords shall be protected from theft and brute-force attacks. This requirement will depend on the built-in hash in Laravel and use bycrypt under the hood.
SS6	Users shall have the right to regain access to their account securely without compromising the account's information. This will depend on the use of Laravel's built-in password rest included in Breeze using a secure email with a token link.
SS7	The users of the platform shall agree to the terms and privacy policy before the creation of their account in the sign-up process.
SS8	The platform shall only collect necessary data that concerns learning, assessment, and system usage.
SS9	The Privacy Policy page of the platform shall be able to explain what data is collected, and why and how it will be used.

4.3.2 Design of Software and/or System and/or Product and/or Processes

4.3.2.1 ER Diagram

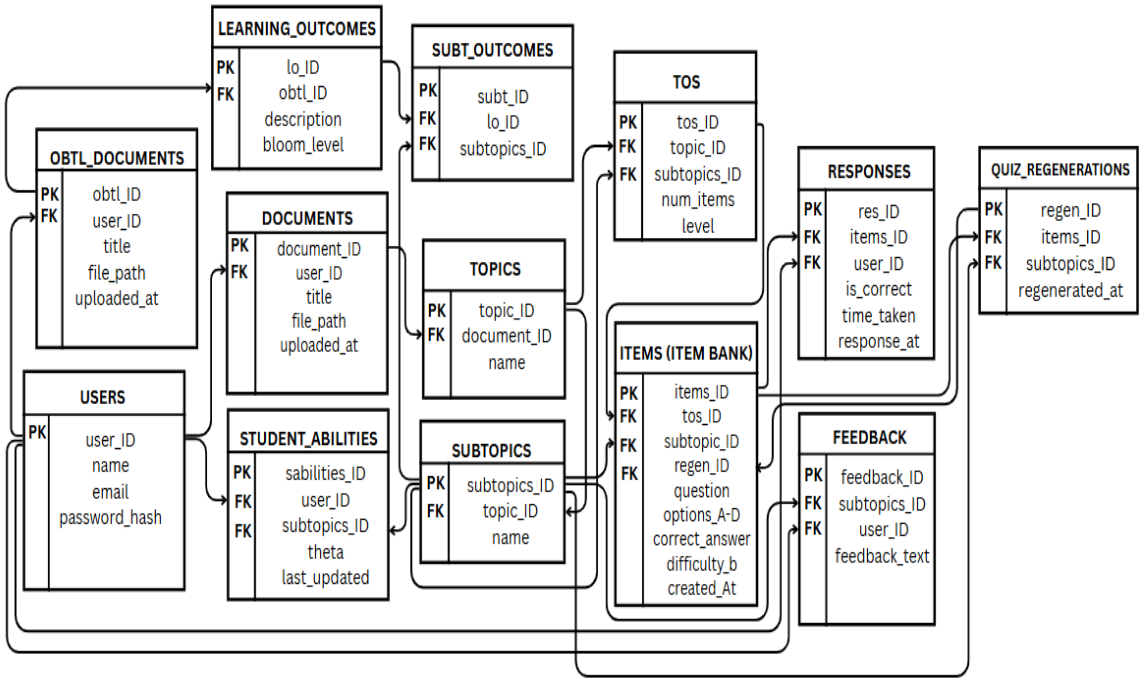


Figure 14. Database Schema of SumakQuiz

Figure 14 shows the Entity Relationship Diagram (ERD) of SumakQuiz, the AI-powered adaptive quizzing platform designed to support personalized learning. The system begins by allowing students to create an account with basic details like their name, email, and password. Once registered, they can upload two types of learning materials: OBTL (Outcome-Based Teaching and Learning) documents and standard lecture files, such as PDFs.

The OBTL documents help define clear learning outcomes and sub-outcomes, each tagged with a specific Bloom's taxonomy level. This structure helps guide the kind of questions the system will generate—mainly focusing on Lower-Order Thinking Skills (LOTS) such as remembering and understanding. Meanwhile, the uploaded PDFs are broken down into topics and subtopics to help organize content and ensure questions are closely tied to specific areas of study.

Using this information, SumakQuiz automatically creates a Table of Specifications (TOS), which outlines how many questions should be generated per topic or subtopic and what difficulty level they should be. The AI then generates quiz questions based on this structure, storing them in an item bank. Each question includes multiple-choice options, the correct answer, its difficulty rating (based on IRT parameters), and the date it was created.

When students take a quiz, their answers are recorded in the system. It tracks whether each answer was correct, how long it took to respond, and when the response was submitted. Over time, the platform builds a learning profile for each student, estimating their ability level (theta) for each subtopic. This helps tailor the difficulty of future quizzes to match the student's progress and maintain an optimal challenge.

The system also supports quiz regeneration which lets students retake quizzes with slightly reworded questions to reinforce learning without simple repetition. These attempts

are logged for tracking purposes. Additionally, SumakQuiz provides personalized feedback per subtopic, giving students insights into their strengths and areas that need improvement.

Altogether, the ERD reflects how SumakQuiz connects AI, content alignment, and adaptive assessment into one integrated platform that supports self-paced, meaningful learning for every student.

4.3.2.2 Functional Decomposition Diagram

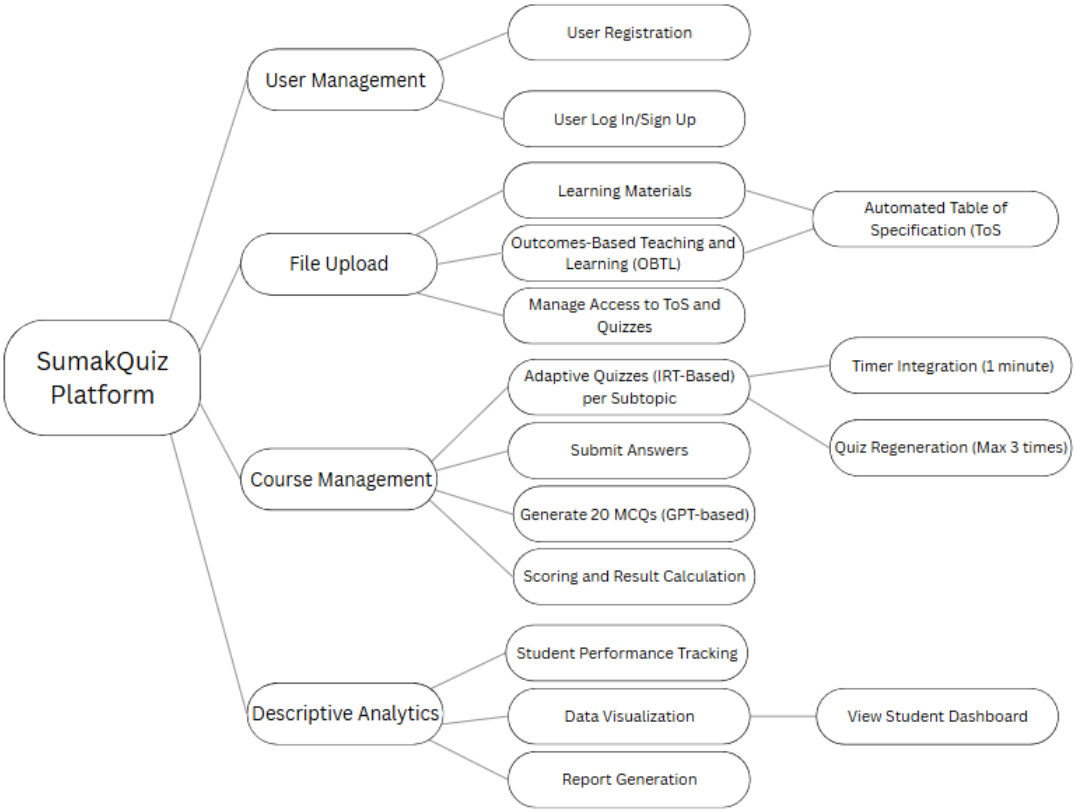


Figure 15. Functional Decomposition Diagram of SumakQuiz

Figure 15 shows the functional decomposition of the SumakQuiz Platform is organized into four main modules: User Management, File Upload, Course Management, and Descriptive Analytics. The User Management module handles user registration and login functionalities to ensure secure access to the platform. The File Upload module allows students to upload lecture materials or Outcomes-Based Teaching and Learning (OBTL) documents, which are used to generate an automated Table of Specification (ToS) limited to LOTS-level cognitive items. The Course Management module facilitates the generation of 20 multiple-choice questions using GPT, delivers adaptive quizzes per subtopic based on the student's estimated ability using the IRT 1PL model, integrates a 1-minute timer per question, and supports quiz regeneration up to three times to provide varied practice while maintaining topic alignment. Finally, the Descriptive Analytics module enables student performance tracking, data visualization, and report generation through a personalized dashboard. This modular structure ensures that SumakQuiz delivers a guided, adaptive, and outcomes-aligned quiz experience tailored to individual learning needs.

4.3.2.3 Operating Environment

This section outlines the technical environment supporting the system, covering both software and hardware specifications for the quiz platform to operate properly. The platform should be compatible with operating systems for it to function efficiently in online conditions. Table 3 lists the key components of the operating environment.

Table 9. Operating Environment

Code	Environment Description
OE1	The platform/system shall run on Windows 7 or higher operating system, Android 10+ or iOS 13+ (mobile).
OE2	Backend services shall run using Python with FastAPI, supported by MySQL and Redis databases.
OE3	AI services (e.g., ToS and quiz generation) shall use OpenAI APIs via Python SDK integration.
OE4	The platform shall be tested in Docker containers for consistency across development and deployment.
OE5	The platform must be compatible with major web browsers such as Chrome, Firefox, and Edge.

4.3.2.4 Design and Implementation Constraints

This section outlines the limitations and conditions that affect the system's development and deployment. These constraints involve technical specifications, programming requirements, hardware specifications, and functional limitations. The table lists the design and implementation constraints considered for this project.

Table 10. Design and Implementation Constraints

Code	Design Constrains and Implementation Constraints Description
DC1	The software shall be programmed using Python, using FastAPI for backend logic and RESTful APIs.
DC2	The system shall apply a simplified 1PL IRT model to adapt question difficulty based on student performance.
DC3	The quiz generation is limited to LOTS-based questions only (Remembering and Understanding).
DC4	The platform supports only English, text-based PDF files. Scanned images and handwritten notes are not supported.

IC1	The development and testing environments will be managed using Docker for consistency.
IC2	AI-generated content shall rely solely on the OpenAI GPT-4 API and internal NLP logic for quiz and feedback generation.
IC3	Platform access requires a stable internet connection for real-time API interactions and quiz delivery.
IC4	The platform shall function on at least the minimum hardware specifications: PC/Laptop- Windows 7+, Intel i3-1155G4 (or equivalent), 4GB RAM, 250GB storage and Mobile devices with at least 4GB RAM, and 64GB storage.

4.3.3 System Development Methodology

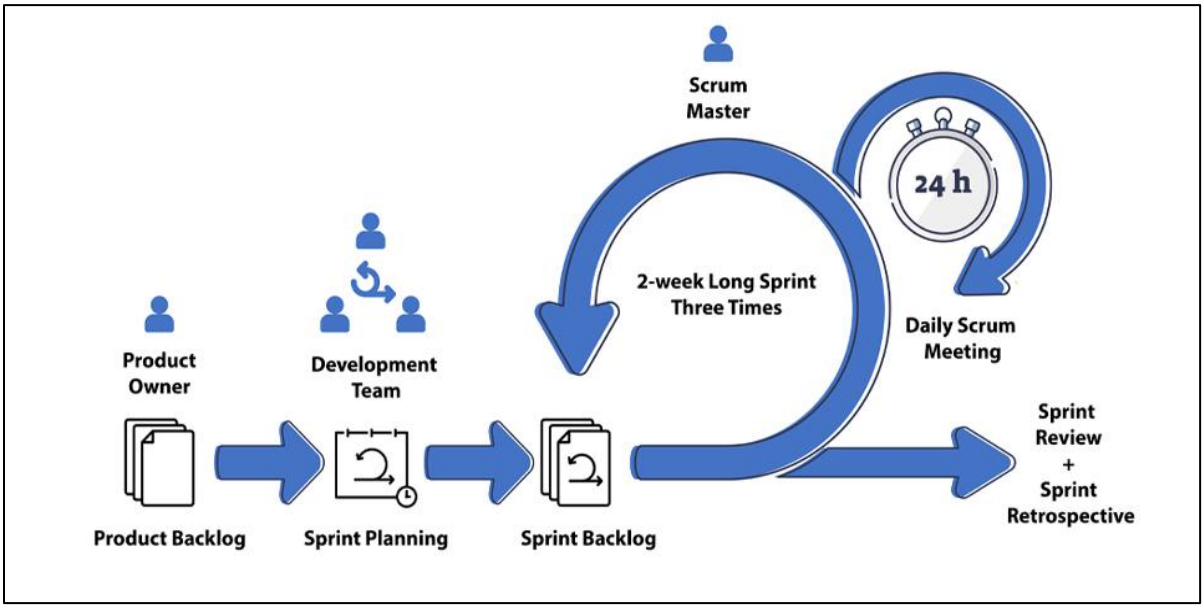


Figure 16. Agile Scrum Model

The development of proposed study SumakQuiz study platform adopted Agile Methodology using scrum framework. This framework will be used as it helps researchers by making the massive project into smaller components making it easier to manage, as well

as guide them to progress with incremental development. This facilitates an iterative approach allowing for an adaptive planning process and continuous feedback from stakeholders. By the adoption of scrum, researchers ensure a time-box process, high-quality standards, and improve productivity.

As the system will be divided into smaller components called sprints which focuses on features and functionalities. This specific methodology possessed three main role such as The Product Owner, representing stakeholders and the one who prioritizes backlog items, The Development Team, which are responsible for the creation of functional increments of the system, and The Scrum Master, the one who guides the team maintaining the sprint schedule and scrum principles.

For SumakQuiz, Mary Gaile Sabado serves as the Product Owner, the Development Team consists of four members they are Clarisa Escoto, John Francis Reanzares, Mary Gail Sabado, and Mark Zablan, while John Francis Reanzares also serve as the Scrum Master, ensuring that the team are aligned with project schedule.

To effectively manage the workload of the system and still maintain its quality, the team followed defined and specific timelines for each task. This structured approach guides the team to deliver the system on time while ensuring its critical functionalities and meeting the expectations of the stakeholder. Moreover, this methodology allows the team to address obstacles, evaluate progress regularly, and integrate feedback to ensure that when the

system reaches the deployment phase it will meet the technical requirements and specifically the user needs.

4.3.3.1 Planning

In this phase, the planning of the proposed study starts with a team meeting, understanding the current study platforms. This allows the researchers to list various features that need to be integrated. These gathered requirements will be done through brainstorming and reviewing literature and studies. Then, researchers prioritize feasibility and importance of these features.

4.3.3.2 Scrum

This framework promotes flexibility, collaboration, and continuous improvement since it allows the development of systems to be broken down into sprints. Daily Stand-ups are also held allowing the research to have a short daily meeting which is 15 minutes long discussing progress, identifying obstacles, and determining what will be the next step. In addition, retrospective, which is a post-sprint meeting, helping researchers to reflect on sprints' success and areas for improvement.

4.3.3.3 Project Schedule

The project schedule included sprints designed to develop, test, and modify SumakQuiz's features. Specific timeframes are allocated for each task to properly divide documentation and development activities, allowing researchers to effectively build a functioning platform.

The schedule offers valuable insights with regards to task completion rates, contingency time, and the project's trajectory, enabling researchers to adapt plans as needed to meet deadlines.

Table 11. Sprint#1: System Design and Initial Setup

TASK	RESPONSIBLE	ESTIMATED START DATE	ESTIMATED END DATE	DURATION IN DAYS
1.1 System Planning	All Members	August 1, 2025	August 4, 2025	4 Days
1.2 Design System Flow	All Members	August 5, 2025	August 8, 2025	4 Days
1.3 Item Bank Setup and Integration	All Members	August 9, 2025	August 15, 2025	7 Days
1.4 Research in OpenAI API	All Members	August 16, 2025	August 20, 2025	5 Days
1.5 UI/UX Design for User Modules	All Members	August 21, 2025	August 27, 2025	7 Days
1.6 Accuracy Testing: Quiz and Table-of-Specification Generation	All Members	August 26, 2025	August 27, 2025	2 Days
1.7 Functionality Testing: Database and Interface	All Members	August 28, 2025	August 30, 2025	3 Days
1.8 Usability Testing: Preliminary Feedback	All Members	August 29, 2025	August 31, 2025	3 Days

Table 11 shows the first sprint of SumakQuiz, focusing on planning how the system works and the way it would look like. This serves as the foundation for the platform as it establishes the creation, structure, flow, and the design of the system. Additionally, in terms of validating the core components of the system accuracy, functionality, and usability testing will be conducted.

Table 12. Sprint #2: Core Functionalities Development

TASK	RESPONSIBLE	ESTIMATED START DATE	ESTIMATED END DATE	DURATION IN DAYS
2.1 Develop User Registration and Login System	All Members	September 1, 2025	September 4, 2025	5 days
2.2 Implement File Upload (PDFs)	All Members	September 5, 2025	September 8, 2025	4 days
2.3 Implement Table-of-Specification and Quiz Generation	All Members	September 9, 2025	September 15, 2025	7 days
2.4 Integrate Adaptive Quiz (per Subtopic)	All Members	September 16, 2025	September 20, 2025	5 days
2.5 Integrate Student Performance Dashboard	All Members	September 21, 2025	September 25, 2025	5 days
2.6 Accuracy Testing: Table-of-	All Members	September 26, 2025	September 27, 2025	2 days

Specification and Quiz Generation				
2.7 Functionality Testing: User Login Systems	All Members	September 28, 2025	September 29, 2025	2 days
2.8 Usability Testing: File Upload	All Members	September 30, 2025	September 30, 2025	1 day

Table 12 focuses on the second sprint which is about Developing Core Functionalities of the system such as user registration and login, file upload, quiz and table-of-specification generation, and descriptive dashboard. Testing during this sprint ensures that these functionalities work well as it will be intended to be a personalized assistant.

Table 13. Sprint #3: Testing

TASK	RESPONSIBLE	ESTIMATED START DATE	ESTIMATED END DATE	DURATION IN DAYS
3.1 Perform System Usability Tests	All Members	October 1, 2025	October 3, 2025	3 days
3.2 Perform System Functionality Tests	All Members	October 4, 2025	October 6, 2025	3 days
3.3 Perform System Accuracy Tests	All Members	October 7, 2025	October 9, 2025	3 days

Table 13 introduces testing such as Usability, Functionality, and Accuracy that will be performed to ensure that the system works perfectly and as expected without the

presence of extensive delays and errors. The goal of this sprint is to preform focusing on efficient testing on core aspects of the system.

Table 14. Sprint #4: Final Refinements and Deployment

TASK	RESPONSIBLE	ESTIMATED START DATE	ESTIMATED END DATE	DURATION IN DAYS
4.1 Refine User Interface Based on Feedback	All Members	October 9, 2025	October 15, 2025	7 days
4.2 Finalize Quiz and Table-of-Specification Generation	All Members	October 16, 2025	October 19, 2025	4 days
4.3 Prepare Deployment	All Members	October 20, 2025	October 21, 2025	2 days
4.4 Conduct First Overall Testing	All Members	October 22, 2025	October 23, 2025	2 days
4.5 Final Testing and Validation	All Members	October 24, 2025	October 25, 2025	2 days

Table 14 will be the final sprint as it focuses on refining the system based on feedback, that ensures the system is fully functional and ready for deployment. In addition, the first overall testing serves a crucial part on this sprint as it will guide the researchers to fully identify remaining issues. On the other hand, final testing and validation will ensure that functional requirements are met.

Table 15. Product Backlog of SumakQuiz

ID	As a User	I want to be able to…	So that…	Priority	Sprint	Status

1	Student	Register an account	I can set up my personal account and gain access to the platform	Must	1	To Do
2	Student	Log in with a registered account	I can access the platform	Must	1	To Do
3	Student	Experience a well-designed interface	I can easily use and understand the system	Must	1	To Do
4	Student	Have a well-planned learning flow	My experience will be intuitive	Must	1	To Do
5	Student	Log out of the platform	I can end my session	Must	1	To Do
6	Student	Upload a lecture and OBTL file	The system can analyze the content and prepare topic-based quizzes	Must	2	To Do
7	Student	Generate Table of Specification	I know the number of items is included per subtopic	Must	2	To Do
8	Student	Take personalized quiz based on my uploaded content	I can test my learning	Must	2	To Do
9	Student	Receive adaptive questions for each subtopic	The difficulty adjusts to my ability	Must	2	To Do
10	Student	See my mastery status per subtopic	I can know which areas to improve	Must	2	To Do
11	Student	Visually track my progress	I can see how I am improving and stay motivated	Must	2	To Do
12	Student	Use all the features as intended through	I can trust the platform	Must	3	To Do

		functionality testing	correctly responds			
13	Student	Use the platform easily through usability testing	I won't be confused while using it	Must	3	To Do
14	Student	Rely on topic-matched and accurate quiz results	I know my quiz results are meaningful	Must	3	To Do
15	Student	Retake quizzes that are regenerated	I can fully grasp the lecture	Must	4	To Do
16	Student	Use an easy-to-use interface	Without confusion, I can navigate the platform	Must	4	To Do

4.3.3.4 Gantt Chart

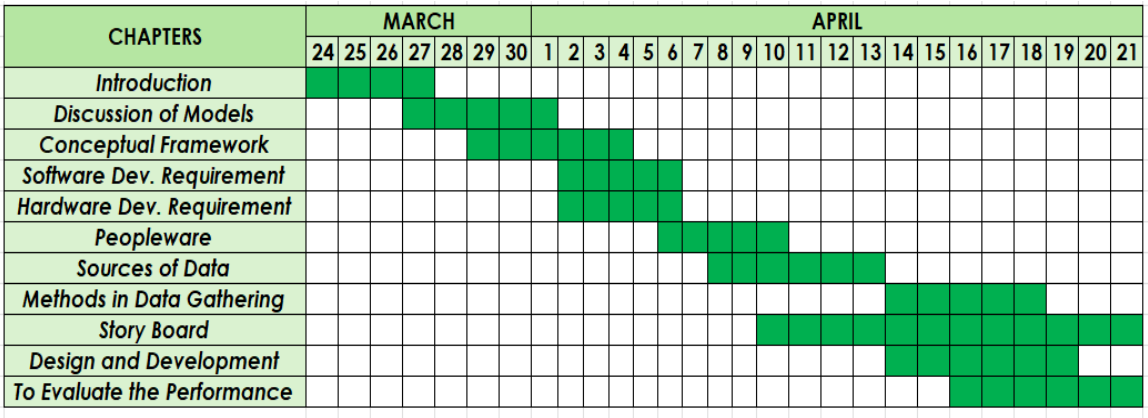


Figure 17. Gantt Chart for Documentation

Figure 26 shows the timeline for completion of each task from chapter 1 to 4 included in the proposal project. This highlights the overlapping schedule to easily manage the documentation process.

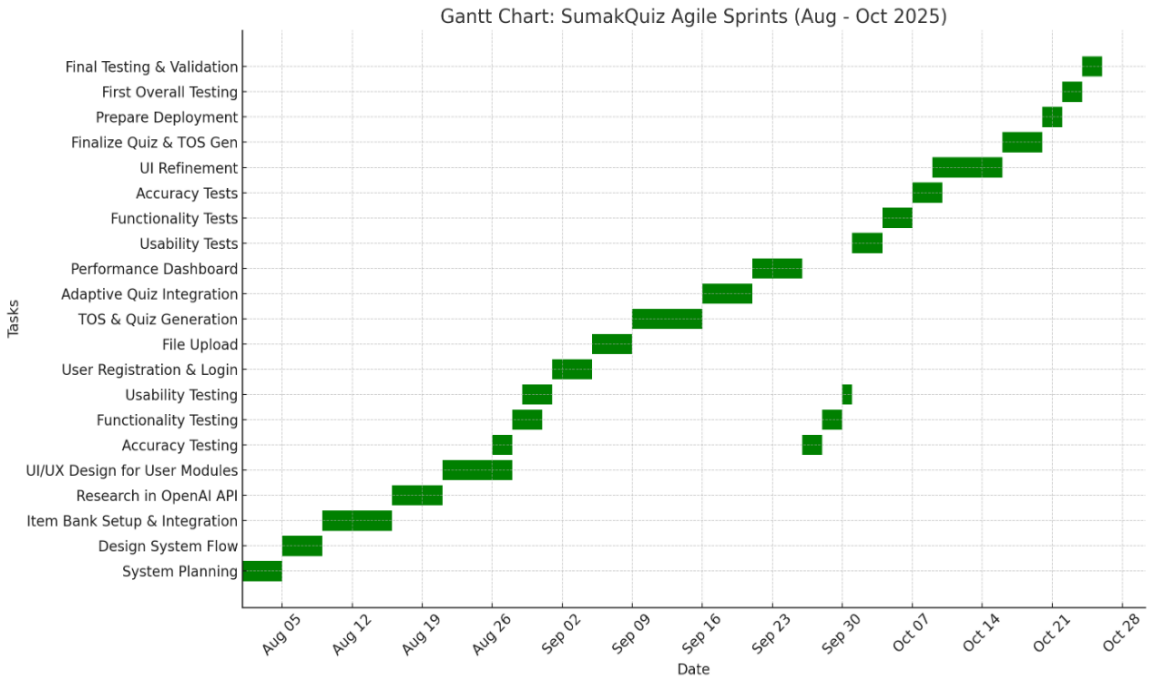


Figure 18. Gantt Chart for Development

Figure 27 shows the timeline of the development for the SumakQuiz System. This schedule is purely based on Agile Scrum methodology, which indicates it is divided into four main sprints such as System Design, Core Development, Testing, and Final Deployment. On each task, there is designated start and end date across August to October 2025.

This timeline will guide the team to ensure the timely delivery of a high-quality system that stays on track and can monitor dependencies.

4.4 To evaluate the performance of SumakQuiz by means of functionality testing, accuracy testing and user experience evaluation.

4.4.1 Functionality Testing (By means of ISO 25010)

This test will be used to validate if the features of the platform function according to its design and specifications. This will determine if the system could reliably support the learning experience of its users. The following table includes the functional test cases:

Table 16. Functionality Test Cases

Registration and Login	
TC1	Verify if a user can register with valid inputs (first name, last name, email, password).
TC2	Verify if a user can log in with registered credentials.
TC3	Verify if a user cannot login for providing an incorrect password.
TC4	Verify if a user cannot register with an invalid email.
TC5	Verify if the user can upload a learning material (PDF) successfully.
TC6	Verify if the system processes the uploaded material and extracts quiz-relevant content.
TC7	Verify that the user can upload or input custom OBTL information.
TC8	Verify that the system generates quiz questions based on the uploaded file and OBTL file/inputs.

TC9	Verify if the difficulty level of generated quizzes adjusts based on user performance.
TC10	Verify if the system generates quizzes that adapt to the user's ability level.
TC11	Verify if the system generates questions with varying difficulty (easy, intermediate, difficult).
TC1	Verify if the quizzes employ interleaving (mixed topic).
TC1	Verify if post-quizzes employ active recall.
TC13	Verify if the system tracks its user's performance.

Table 17. Functionality Test Questions

QUESTIONNAIRE		5	4	3	2	1
Q1	I was able to successfully register with valid information.					
Q2	I could log in using the correct credentials					
Q3	The system prevented login with incorrect credentials.					
Q4	I was not allowed to sign up using an invalid email address.					
Q5	I was able to upload learning materials (PDF) easily.					
Q6	I was able to manually input/upload OBTL information without issues.					
Q7	The system successfully generated a Table of Specification.					
Q8	The system generated quizzes based on the uploaded file and aligned with my OBTL file/inputs.					
Q9	The quiz difficulty changed according to my performance.					
Q10	The platform allowed me to retake quizzes with reworded questions.					
Q11	The platform applied active recall and mixed topics in quizzes.					
Q12	The dashboard accurately tracked and displayed my performance.					

4.4.2 Accuracy Test

To assess the accuracy and instructional quality of AI-generated quiz content, this test will gather feedback from student users regarding the relevance, clarity, and correctness of the questions generated by the platform. Participants will interact with the platform by uploading learning materials, inputting OBTL details, and taking AI-generated quizzes. After completing the quizzes, users will evaluate the quality of the questions through a structured questionnaire. The questionnaire uses a Likert scale to measure aspects such as content accuracy, alignment with learning outcomes, and appropriateness of difficulty.

Additionally, this will also assess the quality of the analytics feature of the platform from the student's perspective through the ISO 25010 standard for Information Quality. This evaluation focuses on the clarity and accuracy of the platform's descriptive analytics which ensures if the elements displayed like graphs and charts are comprehensible for its users and effectively support self-regulated learning. This testing will also verify whether the analytics features such as performance tracking are accurate, understandable, and meaningful. Using this evaluation, this confirms the descriptive analytics is not only accurate but also serves its intended purpose.

Table 18. AI Accuracy Questions

QUESTIONNAIRE		5	4	3	2	1
Q1	The generated questions are factually accurate based on the uploaded learning material.					
Q2	The questions are clearly relevant to the content I uploaded.					
Q3	The questions align with the learning outcome based on the generated ToS.					
Q4	The questions matched the LOTS level (Remember/Understand) of Bloom's Taxonomy.					
Q5	The incorrect answer choices (distractors) are reasonable and not obviously wrong.					
Q6	The questions are clear and grammatically correct.					
Q7	The correct answers are appropriate and supported by the material.					
Q8	The questions difficulty is appropriate for my learning level.					
Q9	The regenerated/reworded questions are non-redundant but maintain the same learning concept.					
Q10	The feedback/recommendation from the system reflects my weak areas.					
Q11	The questions support my learning and knowledge recall effectively.					

Table 19. Descriptive Analytics Accuracy Questions

QUESTIONNAIRE		5	4	3	2	1
Q1	The dashboard shows accurate scores of my quizzes.					
Q2	The platform shows my progress accurately.					
Q3	The displayed graphs and charts are clear and easy to read.					
Q4	The platform's dashboard is organized and easy to follow.					
Q5	The colors, labels, and design of the platform makes it easier for me to understand the graphs and tables.					
Q6	The dashboard made it easier for me to track my progress.					
Q7	The platform provides useful and personalized study recommendations.					
Q7	The dashboard helped me be aware of the topics I am still lacking and need to work on.					

4.4.4 User Evaluation Experience (By means of ISO 25010: Usability Testing)

To understand the user experience with SumakQuiz, this test will gather feedback from the students regarding their experience. The evaluation will focus on user satisfaction, and ease of use of the platform. The participants will utilize the platform and perform common tasks such as signing up, uploading course materials, inputting/uploading OBTL details, taking quizzes, and viewing their performance and recommendations on what topics needed more attention on the dashboard. Following the usage and interaction with the platform, the participants will be asked to complete a questionnaire for evaluation. The questionnaire will consist of quantitative questions which employ a Likert scale to measure aspects such as ease of use and satisfaction. This collected feedback will serve as basis for evaluating the platform's usability and functionality from the user point of view. The findings can be used for future refinements of the platform for better learning experience. The questionnaire for the user evaluation and functionality includes the following questions:

Table 20. User Evaluation Questionnaire

QUESTIONNAIRE	5	4	3	2	1
1. The registration process was easily done by inputting valid information.					
2. I was able to log in by using my registered email and correct password.					

3. I was prevented from logging in when I inputted an incorrect password.					
4. I wasn't able to sign up when I entered an invalid email.					
5. Uploading my learning materials was easy and successful.					
6. I was able to upload an OBTL file.					
7. I was also able to manually enter the OBTL details.					
8. The AI-generated questions reflected the content I uploaded.					
9. The difficulty of the questions adjusted based on how I performed.					
10. I was able to regenerate a quiz for one topic up to three times.					
11. The complexity of the questions given to me varies from easy to difficult.					
12. The quizzes contain mixed different topics.					
13. The system's dashboard allows me to track my learning performance.					

In analyzing Likert scale results, the following range scale and verbal interpretation were used:

Table 21. Likert Scale

Rating	Scale	Verbal Interpretation
5	4.50 – 5.00	Excellent
4	3.50 – 4.49	Very Good
3	2.50 – 3.49	Good
2	1.50 – 2.49	Fair
1	1.00 – 1.49	Poor

Table 21 shows the rating scale and corresponding verbal interpretation used to assess user responses. To analyze the feedback collected through questionnaire, the weighted mean of the responses was computed using the following formula:

Where:

$$\overline{X} = \frac{\sum X}{N}$$

$\bar{x}$ = Weighted Mean

F= Frequency of each rating

x = Numerical value assigned to each rating

N = Total number of respondents

REFERENCES

[1] X. Zhai, X. Chu, C. S. Chai, Y. Li et al., "A Review of Artificial Intelligence (AI) in Education from 2010 to 2020," Complexity, vol. 2021, no. 6, pp. 1–18, Apr. 2021, doi: 10.1155/2021/8812542.

[2] T. N. Fitria, “Artificial Intelligence (AI) in Education: Using AI Tools for Teaching and Learning Process,” in Proc. Seminar Nasional & Call for Paper STIE AAS, Surakarta, Jawa Tengah, Dec. 2021.

[3] K.-J. Laak and J. Aru, "AI and Personalized Learning: Bridging the Gap With Modern Educational Goals," arXiv preprint arXiv:2404.02798, Apr. 3, 2024. doi: 10.48550/arXiv.2404.02798.

[4] S. R. Hirulkar and S. V. Athawale, “Quiz Master AI: An Interactive Machine Learning-Based Quiz Generator,” Int. J. Ingenious Res. Invention Dev. (IJIRID), vol. 3, no. 5, pp. 474, Oct. 2024, doi: 10.5281/zenodo.14208814.

[5] J. Rudolph, S. Tan, and S. Tan, “War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education,” Journal of Applied Learning & Teaching, vol. 6, no. 1, Apr. 2023, doi: 10.37074/jalt.2023.6.1.23.

[6] C. K. Lo, “What is the impact of ChatGPT on education? A rapid review of the literature,” Education Sciences, vol. 13, no. 4, p. 410, 2023, doi: 10.3390/educsci13040410.

[7] M. Chaudhry and E. Kazim, “Artificial Intelligence in Education (AIEd): A high-level academic and industry note 2021,” AI and Ethics, vol. 2, pp. 1–9, 2021, doi: 10.1007/s43681-021-00074-z.

[8] D. Baidoo-Anu and L. Ansah, “Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning,” Journal of AI, vol. 7, 2023, doi: 10.61969/jai.1337500.

[9] X. Wang, S. Wrede, L. van Rijn, and J. Wöhrle, “AI-Based Quiz System for Personalised Learning,” in Proc. 16th Int. Conf. Educ., Res. Innov. (ICERI), Nov. 2023, doi: 10.21125/iceri.2023.1257.

[10] S. Hutt and G. Hieb, “Scaling Up Mastery Learning with Generative AI: Exploring How Generative AI Can Assist in the Generation and Evaluation of Mastery Quiz Questions,” in Proc. 11th ACM, New York, NY, USA: ACM, Jul. 2024, pp. 310–314, doi: 10.1145/3657604.3664699

[11] S. Gupta et al., "Use of an adaptive e-learning platform as a formative assessment tool in the cardiovascular system course component of an MBBS programme," Advances in Medical Education and Practice, vol. 11, pp. 989–996, 2020, doi: 10.2147/AMEP.S267834.

[12] I. Gligorea et al., "Adaptive learning using artificial intelligence in e-learning: A literature review," Education Sciences, vol. 13, no. 12, p. 1216, 2023, doi: 10.3390/educsci13121216.

[13] V. M. Ionescu and M. C. Enescu, "Using ChatGPT for Generating and Evaluating Online Tests," in Proc. 15th Int. Conf. on Electronics, Computers and Artificial Intelligence (ECAI), 2023, pp. 1–6, doi: 10.1109/ECAI58194.2023.10193995.

[14] J. Firth, I. Rivers, and J. Boyle, "A systematic review of interleaving as a concept learning strategy," Review of Education, vol. 9, no. 1, pp. 1–32, 2021, doi: 10.1002/rev3.3266.

[15] S. Heitmann et al., "Adaptive practice quizzing in a university lecture: A pre-registered field experiment," Frontiers in Psychology, vol. 12, 2021.

[16] L. Lim, S. H. Lim, and W. Y. R. Lim, "Efficacy of an adaptive learning system on course scores," Education and Information Technologies, vol. 27, pp. 313–329, 2020.

[17] Y. Haryanto et al., "Item selection on the Moodle-based computerized adaptive test," International Journal of Interactive Mobile Technologies, vol. 14, no. 6, pp. 125–132, 2020.

[18] J. Xu et al., "Active recall strategies associated with academic achievement in young adults: A systematic review," Journal of Educational Psychology, vol. 113, no. 5, pp. 870–888, 2021.

[19] N. Wengrowicz et al., "Modeling with real-time informative feedback: Implementing and evaluating a new massive open online course component," J. Sci. Educ. Technol., vol. 32, no. 6, pp. 884–897, 2023.

[20] S. Liu and D. Lin, "Developing and validating an analytic rating scale for a paraphrase task," Assessing Writing, vol. 53, p. 100646, 2022.

[21] R. Setyorini, "The dominance of LOTS in summative assessment: The challenge of improving HOTS in Indonesian language learning," Canadian Journal of Educational and Social Studies, vol. 5, no. 1, pp. 114–125, 2025, doi: 10.53103/cjess.v5i1.295.

[22] H. Lee and N. Bosch, "Subtopic-specific heterogeneity in computer-based learning behaviors," Int. J. STEM Educ., vol. 11, no. 1, p. 61, 2024.

[23] J. Hua and R. Nagappan, "Research on the influencing factors of students' autonomous learning ability in higher vocational colleges and universities in the context of artificial intelligence," Educational Review (USA), vol. 8, no. 9, 2024.

[24] J. S. Barrot, E. M. Masangya, and J. I. G. Lira, "Outcomes-based teaching and learning assessment instrument for teachers in higher education," Curriculum and Teaching, vol. 38, no. 2, pp. 21–35, Nov. 2023, doi: 10.7459/ct/38.2.03.

[25] B. R. B. Milloria, A. M. D. Marzon, and L. M. C. Derasin, "Investigating AI-integrated instruction in improving academic performance of senior high school students in the Philippines," Journal of Harbin Engineering University, vol. 45, no. 6, 2024.