The International Society of Data Scientists
Journal of Data Science and Artificial Intelligence
Vol. 4, No. 1. Online ISSN: 2831-4794

# An Intelligent Platform for Assessing Program Learning Outcomes: Leveraging Data Processing and AI-Driven Reporting

Duc Dat Pham[a,b], Mai Thanh Nguyen Quynh[a,b], Mai Oanh Nguyen Ngoc[b,c], Tan Duy Le[*a,b], Kha Tu Huynh[a,b]

[a]International University, Ho Chi Minh City, Vietnam; [b]Vietnam National University, Ho Chi Minh City, 700000, Vietnam; [c]HCMC University of Technology, VNU-HCM

## ABSTRACT

The EVALLOS platform introduces a structured and data-driven approach to the assessment of Program Learning Outcomes (PLOs) and Course Learning Outcomes (CLOs) in higher education. Leveraging advanced data processing, artificial intelligence (AI), and dynamic visualization tools, the system provides a unified framework for evaluating educational outcomes. EVALLOS automates key processes, including data collection, CLO-PLO mapping, and report generation, ensuring precision, efficiency, and scalability. With features such as real-time dashboards, AI-generated insights, and actionable recommendations, the platform empowers educators to refine teaching strategies, address learning gaps, and align curriculum with institutional goals. Additionally, EVALLOS facilitates accreditation compliance by delivering comprehensive reports that integrate quantitative metrics with qualitative analysis. EVALLOS supports institutional decision-making, continuous improvement, and the effective application of AI in education.

**Keywords:** Program Learning Outcomes (PLOs), Course Learning Outcomes (CLOs), Educational Assessment, Data Visualization, Accreditation, Learning Analytics, Higher Education.

## 1. INTRODUCTION

In the contemporary landscape of higher education, the assessment of Program Learning Outcomes (PLOs) has emerged as a pivotal component in ensuring the quality and effectiveness of educational programs. PLOs articulate the specific knowledge, skills, and competencies that students are expected to acquire upon completing a program, serving as essential benchmarks for curriculum design and instructional effectiveness [1]. As institutions strive to meet accreditation standards and enhance student learning experiences, the need for robust assessment frameworks that accurately measure PLO attainment has become increasingly critical [2].

Traditional methods of evaluating learning outcomes frequently depend on manual data capture and analysis, which can be labor-intensive and susceptible to human error. The aggregation of student performance data from a variety of sources, including examinations, assignments, and surveys, is typical in these conventional approaches. This is followed by a time-consuming process of analysis and reporting. In addition to impeding the ability of educators to effectively identify trends and areas for improvement, these methods may also fail to provide a comprehensive view of student performance, which may impede timely decision-making [3]. Additionally, the absence of real-time data can impede the capacity of institutions to timely address emergent educational challenges, thereby impacting the overall quality of education provided.

The educational assessment landscape is substantially transformed by technology, which employs sophisticated data management, advanced analytics, and the deployment of artificial intelligence (AI), particularly Large Language Models (LLMs) such as GPT-3.5 and LLama3.

Data Processing and Calculation: The meticulous administration and processing of data is the foundation of the integration of technology in educational assessments. This phase entails the acquisition of a diverse array of educational data points, including student responses and performance metrics. These data points are subsequently processed and analyzed to calculate comprehensive scores and other pertinent performance indicators. Technology surpasses the constraints of conventional methods by guaranteeing the precise and efficient management of large and intricate datasets through these sophisticated data processing capabilities [4], [5].

Report Generation and Visualization: Following the data processing phase, LLMs such as LLama3 are employed to produce comprehensive reports and facilitate advanced data visualization. These models employ the processed data to generate comprehensive reports that provide profound insights into student performance, identifying areas for refinement and strengths. Furthermore, AI-driven visualization tools visually represent this data, rendering it more comprehensible

and accessible to both students and educators. These instruments enable a more rapid and comprehensive understanding of educational outcomes, which in turn improves educational planning and decision-making [6], [7].

Through these enhancements, technology not only streamlines the evaluation process but also adds layers of precision and personalization. It supports a more data-driven approach to education, helping institutions develop more responsive teaching strategies and personalized learning experiences.

The PLOs assessment platform is intended to offer a contemporary method for evaluating and assessing student learning outcomes while significantly aiding the educational quality accreditation process. The amalgamation of data processing technology and AI-driven reporting enables educational institutions to accurately evaluate student progress relative to established educational objectives while also furnishing substantial proof for accreditation organizations.

The platform (EVALLOS) utilizes automated analysis to discover patterns, strengths, and weaknesses in learning results, therefore recommending solutions to enhance educational programs. This is particularly crucial when universities and other educational institutions must exhibit adherence to the educational and quality standards mandated by accreditation bodies. Consequently, the platform serves as both an efficient assessment instrument and a crucial resource for sustaining and enhancing the quality of university education.

## 2. BACKGROUND AND RELATED WORKS

### 2.1 Overview of PLO Assessment Methods

PLOs are critical elements in higher education, delineating the information, skills, and competencies that students are anticipated to acquire before graduation. They direct curriculum development, instructional strategies, and evaluations, ensuring that educational programs conform to academic standards and professional criteria. PLOs delineate the objectives of a program and function as essential metrics for assessing educational efficacy and upholding accrediting criteria. Their stringent evaluation enables institutions to exhibit their dedication to quality education and ongoing enhancement, fulfilling the demands of accrediting organizations, students, and employers simultaneously [8].
In the domain of accreditation and quality assurance, PLOs are essential. Accrediting bodies employ these outcomes to assess whether educational programs conform to specified quality criteria and effectively provide the desired educational value. This evaluation process entails examining the clarity and relevance of PLOs, the strategies employed to attain them, and their effectiveness in preparing students for success after graduation. This comprehensive evaluation urges institutions to undertake continuous self-assessment and improvement of their educational processes to align with the changing demands of the workforce and society. Moreover, clearly articulated and accomplished PLOs augment an institution's credibility and appeal, so enhancing its capacity to secure funds, recruit distinguished staff, and attract exceptional students.

Assessments of PLOs efficiently utilize Bloom's Taxonomy to guarantee that educational programs thoroughly encompass a wide range of cognitive talents. This taxonomy classifies cognitive capacities into a hierarchical progression from fundamental knowledge acquisition to sophisticated problem-solving skills. At the fundamental levels, "Remembering and Understanding," students are evaluated on their capacity to retrieve facts and fundamental concepts, as well as to understand material accurately. At the "Applying and Analyzing" stage, assessments necessitate that students exhibit their capacity to use knowledge in novel contexts and to analyze arguments or facts, discerning fundamental patterns and relationships. At the pinnacle of this taxonomy are the "Evaluating and Creating" levels, wherein students are required to critically assess material and synthesize novel, logical concepts, demonstrating their utmost mastery and comprehension. These organized cognitive categories provide examinations that focus on particular levels of comprehension, guaranteeing a comprehensive evaluation of the depth and scope of student learning, essential for attaining the explicitly established PLOs [9].

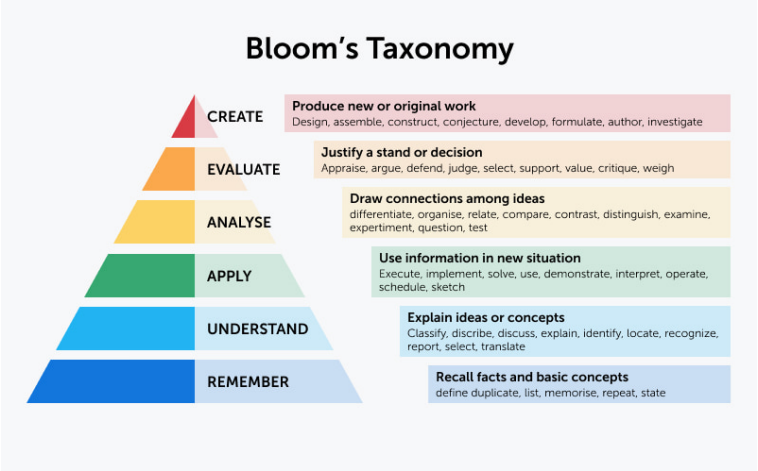



Figure 1. The revised blooms taxonomy (Source: Bloom's Taxonomy) [9] .

### 2.2 Detailed Methodology for Calculating PLOs

The methodology for calculating PLOs relies on aggregating data from CLOs achieved across various levels, from individual exam teams to classes, courses, and ultimately to program-level assessments. This process ensures that PLOs reflect a comprehensive evaluation of student performance and alignment with program goals.

The calculation of CLOs [10] at the exam team level involves a systematic approach to evaluate student performance based on specific assessments. This process starts with collecting scores for individual questions or rubric criteria from students within an exam team. Each question or rubric is mapped to one or more CLOs using a predefined matrix that links assessment tasks to learning objectives. To determine achievement, students' scores are compared against a minimum pass threshold, typically set at 50%.

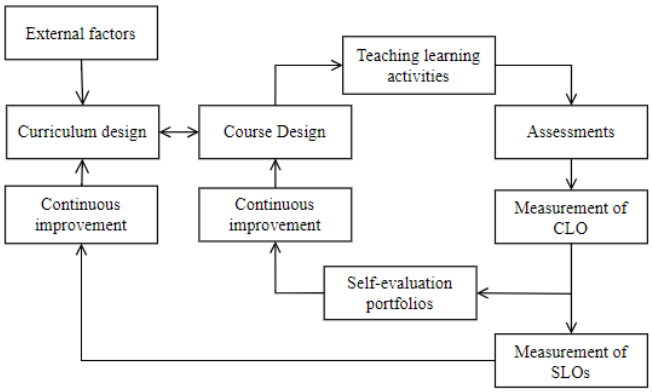



Figure 2. Workflow for calculating CLOs achievement by exam teams in EVALLOS.

The number of students passing each question is aggregated, and the average CLO score is calculated using the formula:

$$\text{Average CLOs} = \frac{\sum_{i=1}^{n} \text{Number of Students Passing Question } i}{n}$$

This score is then normalized to a percentage, reflecting the overall CLO achievement for the exam team:

$$\text{CLOs Achievement} = \frac{\text{Average CLOs} \times 100}{\text{Total Students in Exam Team}}$$

The results provide a clear picture of how well students in the exam team meet the expected learning outcomes. A high percentage indicates strong alignment with learning objectives, while lower scores highlight areas needing instructional or assessment adjustments. This structured process ensures that evaluations are transparent, consistent, and aligned with educational goals.

Once the CLO achievements for each exam team are calculated, these results are aggregated to determine the overall CLO achievement at the class level. This aggregation ensures a holistic evaluation of student performance across all assessments conducted by different exam teams within the class. The formula used for this calculation is:

$$\text{CLO Achievement (Class)} = \sum_{i=1}^{n} \left(\text{CLO Achievement (Assessment}_{i)} \times \text{Weight}_i\right)$$

Here,

- $CLOAchievement(Assessment(_i))$ represents the CLO achievement calculated for the $i$-th exam team.
- $Weight(_i)$ is its assigned weight.

This formula ensures that the contributions of all exam teams are appropriately weighted, reflecting their relative significance in the class-level CLO achievements. By incorporating this weighted approach, educators can account for the varying impacts of different assessments on the learning outcomes, providing a more accurate and balanced measurement of the class's overall performance.

These class-level CLO achievements are further aggregated across all classes in a course to calculate the course-level CLO achievement. This step provides a comprehensive view of how well students are achieving the learning outcomes across the entire course. The formula for calculating course-level CLO achievement is:

$$\text{CLO Achievement (Course)} = \frac{\sum_{i=1}^{n} \text{CLO Achievement (Class}_{i)}}{n}$$

Here,

- $CLOAchievement(Class(_i))$ represents the achievement of a specific CLO in the $(i) - th$ class of the course.
- $(n)$ is the total number of classes associated with the course.

This layered approach ensures that performance is assessed systematically from individual exam teams to classes and then to the overall course. It provides educators with actionable insights into how learning outcomes are being achieved at every level, supporting data-driven decisions for improving teaching and learning strategies.

The relationship between CLOs and PLOs is established through a structured mapping matrix. This matrix defines how each CLO contributes to one or more PLOs, ensuring alignment between course-specific objectives and broader program-level goals. For instance, as illustrated in the matrix below, CLO1 contributes to both PLO1 and PLO4, while CLO2 supports PLO2 and PLO3. Each "X" in the matrix represents a direct contribution of a CLO to a particular PLO.

Table 1. Mapping table between CLO and PLO

| **PLOs and CLOs Matrix** | | | | |
|---|---|---|---|---|
| **CLOs** | **PLO1** | **PLO2** | **PLO3** | **PLO4** |
| CLO1 | X | | | X |
| CLO2 | | X | X | |
| CLO3 | X | | | |
| CLOn | | X | X | X |

To reflect the varying importance of each CLO in achieving specific PLOs, weighting factors can be assigned. This ensures that the contribution of each CLO is proportional to its significance in attaining the corresponding PLO. Such a mapping system provides a clear and systematic framework for assessing the alignment between course outcomes and program goals, enabling institutions to monitor and improve the effectiveness of their educational programs.

The calculation of PLOs [11] involves evaluating the extent to which students meet the defined learning objectives by aggregating their performance across various courses and learning activities. The first step is to compute the average PLO score, which is determined by summing the number of students passing the associated CLOs and dividing it by the total number of CLOs being assessed. This can be expressed using the formula:

$$\text{AveragePLOs} = \frac{\sum_{i=1}^{n} \text{No. of Students Passing CLOs}_i}{n}$$

Once the average PLO score is determined, it is converted into a percentage to reflect the overall achievement of the PLO. This percentage is calculated by normalizing the average PLO score to the total number of students in the program, using the formula:

$$\text{PLOs Achievement} = \frac{\text{AveragePLOs} \times 100}{\text{Total Students in Program}}$$

These calculations provide a standardized and transparent method for assessing how well the program-level outcomes are being met, ensuring that the results are consistent and actionable for program evaluation and improvement.

### 2.3 Challenges in Traditional Assessment Methods

Traditional assessment frameworks frequently depend on manual data processing tools, such as Excel, to calculate PLOs. While these tools offer familiarity and accessibility, they have significant limitations in terms of scalability and efficiency. Manual processing is time-consuming, particularly when handling large datasets, and often prone to human error. These inaccuracies can lead to compromised assessment reliability, impacting both immediate feedback and long-term program planning. Additionally, the need for repetitive data entry and analysis increases workload for educators and administrators, detracting from other critical teaching and program improvement activities. According to recent studies, manual methods are insufficient for managing the complexities of modern educational datasets and fail to provide the timely, precise insights required for informed decision-making [11].

The increasing volume and complexity of educational data necessitate a shift toward automated and advanced data handling methodologies. Automated systems, supported by robust data processing algorithms, address the limitations of traditional methods by improving efficiency and reducing the risk of errors. These systems enable real-time analysis, providing accurate, actionable insights into PLO attainment and overall program performance. Furthermore, sophisticated platforms integrate various data sources, streamlining the evaluation process while offering educators a comprehensive view of student achievement across multiple dimensions. Research highlights that such tools not only improve the reliability of assessment results but also empower educators to implement targeted interventions, enhancing the overall quality of education [12].

### 2.4 Review of Existing AI-Driven Platforms

Artificial Intelligence has profoundly altered educational evaluations through task automation, increased efficiency, and the enhancement of data-driven decision-making. Numerous current systems demonstrate the varied applications of AI in education. Gradescope use machine learning to automate the grading of handwritten, typed, and multiple-choice assignments, therefore substantially decreasing manual labor while ensuring precision. This platform aids educators in recognizing prevalent errors in student submissions, hence enabling focused criticism and enhanced comprehension of learning deficiencies [13].

Likewise, the utilization of language models such as OpenAI's GPT-3 has become prevalent in educational technologies for producing student feedback reports, summarizing performance trends, and developing instructional content. The advanced natural language processing skills of GPT-3 facilitate the production of contextually precise and coherent textual insights. Nevertheless, whereas these systems are proficient in generating individual feedback, they frequently exhibit insufficient integration with overarching program-level evaluations [14].

Platforms such as DreamBox and Smart Sparrow emphasize adaptive learning, employing AI to provide customized content according to each student's learning velocity and preferences. These systems adapt instructional content in real-time according to performance data, improving engagement and personalized learning results. Nonetheless, these platforms are confined to micro-level interventions and fail to address broader programmatic objectives, such as assessing PLOs or facilitating institutional accreditation [15].

Notwithstanding their strengths, current AI-driven platforms frequently demonstrate disjointed functionalities. Some individuals excel in automating assessments, while others concentrate on personalized education or producing textual analyses. A deficiency persists in systems that effectively amalgamate data from diverse sources, facilitate multi-tiered analysis, and endorse thorough program-level evaluations.

The suggested platform - EVALLOS rectifies these deficiencies by implementing a comprehensive solution that amalgamates robust data processing, sophisticated analytics, and AI-driven reporting utilizing the Llama3 model. This system employs a comprehensive approach to educational assessments, in contrast to existing platforms that concentrate on single capabilities. The Llama3 approach improves natural language generation for producing comprehensive and contextually relevant reports for schools, administrators, and accreditation organizations.

This platform's principal attribute is its organized data pipeline, which automates the acquisition, preprocessing, and amalgamation of data from many sources, such as examination results, rubrics, and surveys. This thorough method allows the system to accurately compute metrics such as CLOs and PLOs. CLO accomplishments are calculated for each

examination team and consolidated at the class and course levels prior to their alignment with PLOs, offering a comprehensive perspective on student and program performance.

Furthermore, the platform employs sophisticated data visualization methods, converting intricate datasets into user-friendly dashboards and visual reports. These visualizations allow instructors and administrators to swiftly discern trends, including opportunities for enhancement in specific courses or strengths in particular program results. This degree of knowledge enables evidence-based decision-making, essential for ongoing enhancement and quality assurance.

A notable feature of the platform is its capacity to dynamically align CLOs with PLOs, enabling institutions to monitor the contribution of specific course outcomes to overarching program objectives. This computerized mapping guarantees compliance with accrediting standards and diminishes the manual labor associated with conventional assessment techniques. The incorporation of weighting elements allows the system to represent the relative significance of different examinations, offering a refined comprehension of accomplishment measures.

In contrast to platforms such as Gradescope, GPT-3, and adaptive learning systems, the suggested platform provides an integrated framework that amalgamates data management, AI-powered analytics, and pragmatic reporting. By addressing both micro- and macro-level assessment requirements, it enables educators to elevate teaching quality, enhance learning outcomes, and facilitate institutional accreditation initiatives. This thorough approach positions the platform as a strong answer for contemporary educational difficulties [13], [14].

## 3. METHODOLOGY

The methodology employs advanced technologies, particularly large language models (LLMs), to automate the generation of reports and data visualizations for evaluating CLOs and PLOs. This strategy converts raw data into actionable insights by generating comprehensive reports and dynamic visualizations that effectively convey student performance trends, underscore strengths, and pinpoint areas needing enhancement. By merging precise data analysis with intuitive visual representations, the system enables educators to refine teaching methodologies and aids institutional leaders in making informed, evidence-based decisions. This integration not only improves transparency and efficiency in assessing learning outcomes but also guarantees ongoing enhancement in teaching quality and alignment with accreditation standards, ultimately augmenting the overall effectiveness of educational programs.

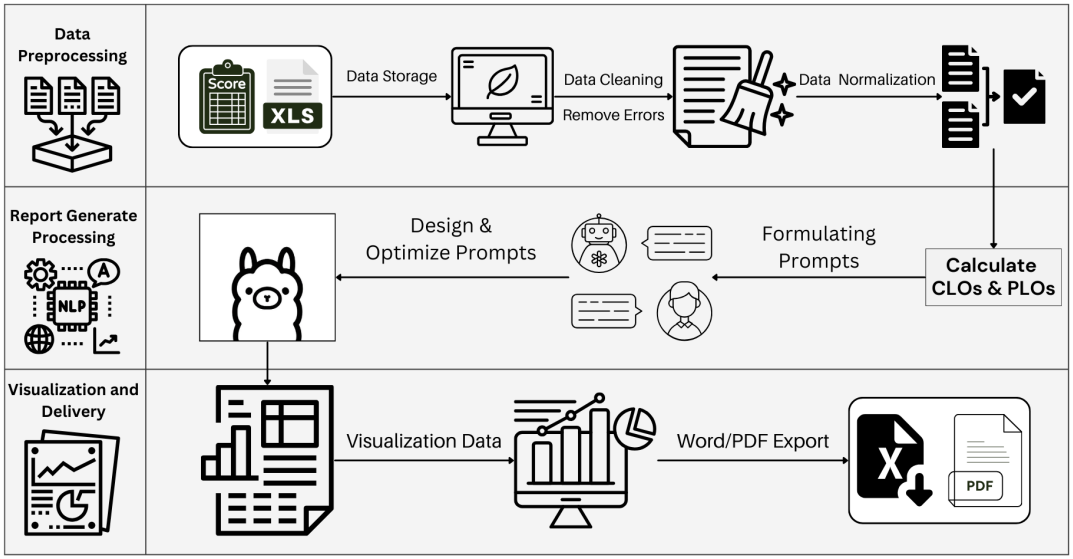



Figure 3. AI-Driven Workflow for Generating PLO and CLO Reports

### 3.1 Data Collection and Storage

The EVALLOS platform employs a systematic methodology for data collection, utilizing a MongoDB database to store and organize a variety of datasets, including assessment results, test scores, assignment grades, attendance records, and qualitative feedback from lecturers. MongoDB [16], a NoSQL database, is selected for its capacity to manage unstructured data, providing scalability and flexibility for large and diverse datasets. This database architecture facilitates seamless integration with Learning Management Systems (LMS) and other institutional data sources, promoting efficient efficiently manages large and diverse datasets, ensuring scalable and fast data retrieval.

By leveraging MongoDB's document-based model, EVALLOS streamlines the representation of intricate educational records. The platform accommodates data points such as grade structures, course mappings, and rubrics, ensuring that all inputs are structured and accessible for subsequent processing stages.

### 3.2 Data Preprocessing

Following data collection, the preprocessing phase guarantees that datasets are clean, consistent, and ready for analysis through a series of automated procedures. Initially, data cleaning detects and corrects errors such as absent scores, duplicate entries, or inconsistent formatting, ensuring that only valid, high-quality data advances to subsequent phases. Subsequently, normalization standardizes scores from diverse assessments to a uniform percentage scale, facilitating consistent processing across courses, exam types, and institutions. Ultimately, the platform aligns assessment data with specific CLOs utilizing a predefined mapping matrix. This alignment ensures clarity and traceability by directly connecting each assessment to its intended learning objectives. These preprocessing steps are efficiently automated through custom Python scripts and MongoDB's aggregation framework, minimizing the risk of human error and optimizing the data preparation process.

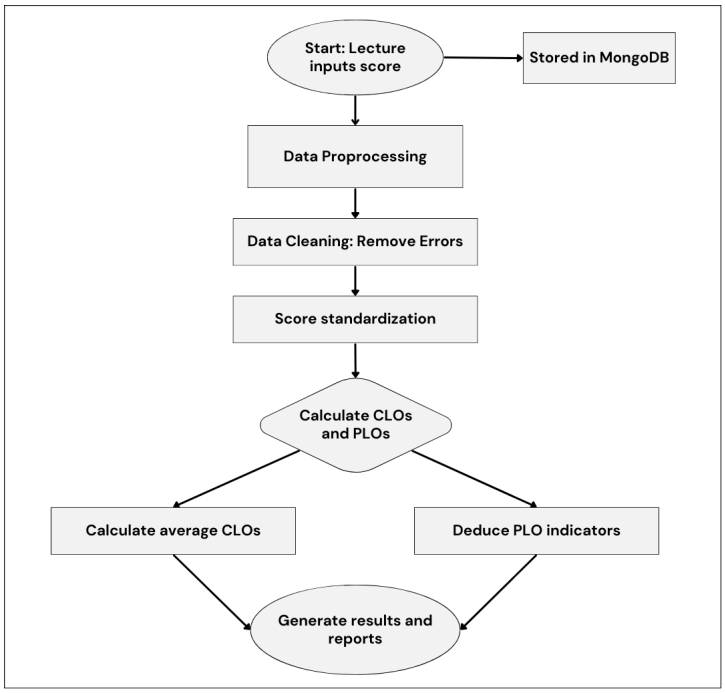



Figure 4. Step by step for data preparation and retrieval

### 3.3 AI-Driven Processing of CLOs and PLOs

EVALLOS leverages artificial intelligence and advanced algorithms to analyze and evaluate both CLOs and PLOs through a multi-level computational framework. For CLOs, the platform calculates achievement levels by assessing student performance against predefined success thresholds, with each assessment weighted according to its significance within the course structure. The system automatically aggregates CLO performance data across exam teams, classes, and entire courses, providing educators with a holistic view of student learning outcomes.

Building on these CLO results, PLO achievements [10] are determined by mapping CLOs to their corresponding PLOs, creating a transparent and hierarchical evaluation framework. This approach ensures that program-level outcomes are directly linked to individual course-level assessments. By employing weighted averages, the platform accurately represents the varying significance of different CLOs in contributing to specific PLOs, delivering actionable insights for continuous improvement.

### 3.4 AI-Driven Processing of CLOs and PLOs

EVALLOS is characterized by its incorporation of a self-hosted large language model (LLM), Llama3 [17], which produces comprehensive and actionable data for educators and administrators. These reports integrate data insights, trends, and actionable recommendations into a coherent and accessible style, facilitating stakeholders in making informed decisions regarding student performance and program results. Llama3 enables detailed, context-aware analysis tailored to educational requirements [18].

EVALLOS utilizes sophisticated quick engineering methodologies to direct the LLM in producing precise, pertinent, and high-caliber results. The prompts are designed to produce outputs tailored to institutional needs and user expectations, improving the relevance of the generated reports. This meticulous optimization procedure improves the LLM's capacity to integrate intricate datasets into coherent narratives and visual representations.

The platform's computing skills are enhanced by a strong hardware infrastructure. Llama3 is implemented on high-performance servers with NVIDIA A100 GPUs and AMD EPYC CPUs, facilitating efficient large-scale data processing and swift report production. This sophisticated configuration allows EVALLOS to scale seamlessly, catering to institutions of diverse sizes and complexities. The platform's integration of hardware and software guarantees stability, scalability, and performance, whether for handling data for an individual course or assessing outcomes across an entire university. Collectively, these attributes render EVALLOS a revolutionary instrument for enhancing the quality and efficacy of educational evaluations.

### 3.5 Integrative Visualization and Recommendations for Enhancing Educational Outcomes

EVALLOS improves educational assessment with advanced visualization features and practical recommendations, allowing institutions to implement a data-driven strategy for ongoing enhancement. The platform utilizes sophisticated visualization capabilities driven by D3.js to display intricate data in interactive and comprehensible formats, including bar charts, pie charts, and heatmaps. These interactive visualizations enable educators and administrators to examine performance trends, pinpoint areas for enhancement, and get actionable information. Real-time dashboards offer consolidated data across many levels, from specific CLOs to comprehensive PLOs, enabling stakeholders to swiftly identify underperforming cohorts or deficient courses.

EVALLOS not only visualizes data but also produces professional-grade reports in formats such as PDF and Word. These reports integrate comprehensive narratives, visual summaries, and customized suggestions, enabling stakeholders to convey findings effectively and make educated decisions. AI-generated recommendations are essential to the platform, providing targeted actions such as correcting ineffective exams, adjusting teaching tactics to bridge learning gaps, and realigning course objectives with overarching program-level outcomes. These recommendations are based on trend assessments and aim to improve alignment with institutional objectives.

Additionally, EVALLOS integrates user feedback loops, facilitating the ongoing enhancement of its algorithms and results. EVALLOS equips institutions with a comprehensive toolkit for understanding and enhancing educational results through the integration of sophisticated visualization, data analysis, and practical recommendations, hence ensuring adaptability to the evolving issues in the educational landscape.

## 4. AI-DRIVEN REPORTING AND VISUALIZATION: TRANSFORMING EDUCATIONAL INSIGHTS

The EVALLOS platform utilizes sophisticated artificial intelligence and data visualization tools to transform the evaluation of educational results. The integration of AI-driven reporting with dynamic visualization improves the usability and interpretation of complicated data, enabling educators and administrators to make educated decisions. This novel methodology directly confronts the limitations of conventional data analysis techniques, yielding explicit, actionable insights that enhance decision-making inside educational environments.

EVALLOS produces comprehensive reports tailored to meet certain educational requirements, providing two main categories of information. PLOs Reports offer a detailed summary of PLO accomplishments within academic programs. These reports encompass essential information, including the proportion of students achieving PLO benchmarks and longitudinal trend analysis across several semesters. These insights provide administrators with a comprehensive view of program effectiveness, allowing them to pinpoint areas for enhancement, revise strategic initiatives, and ensure compliance with accreditation standards.

CLOs Reports provide detailed analysis of CLO performance at the subject, class, and examination team levels. These reports evaluate pass rates, performance trends, and instructional efficacy, assisting educators in identifying learning deficiencies, enhancing teaching strategies, and realigning course objectives to more effectively achieve program-level outcomes. EVALLOS combines quantitative data with contextual narratives to provide a balanced viewpoint that enables institutions to improve teaching quality, optimize assessments, and effectively attain institutional goals.

EVALLOS employs advanced visualization capabilities with D3.js to improve the accessibility and comprehension of intricate data. These visualizations, comprising bar charts, heatmaps, and pie charts, effectively illustrate performance trends and emphasize critical information. Stakeholders can engage with real-time dashboards, examining aggregated data across multiple levels—from individual CLOs to comprehensive PLO performance. This interactivity enables users to swiftly discover underperforming cohorts, low-achieving courses, or systemic tendencies, facilitating timely and informed interventions.

The platform offers professional report formatting in PDF and Word, enabling effortless sharing and accessibility. These reports have intuitive formats that integrate comprehensive narratives with visual summaries to facilitate rapid understanding and informed decision-making.

A notable aspect of EVALLOS is the incorporation of AI-generated recommendations in its reports, allowing institutions to adopt a proactive stance towards ongoing enhancement. The system analyzes trends and performance data to produce actionable recommendations tailored to specific requirements, including revising underperforming assessments or individual questions, adjusting teaching strategies to address identified learning deficiencies, and realigning course objectives with program-level outcomes to improve curriculum coherence. These focused, data-driven recommendations guarantee that actions are prompt and efficacious, precisely tackling areas of concern. EVALLOS features an iterative feedback mechanism enabling stakeholders to enhance algorithms and prompts, hence always refining the system's outputs to meet changing educational problems and institutional requirements.

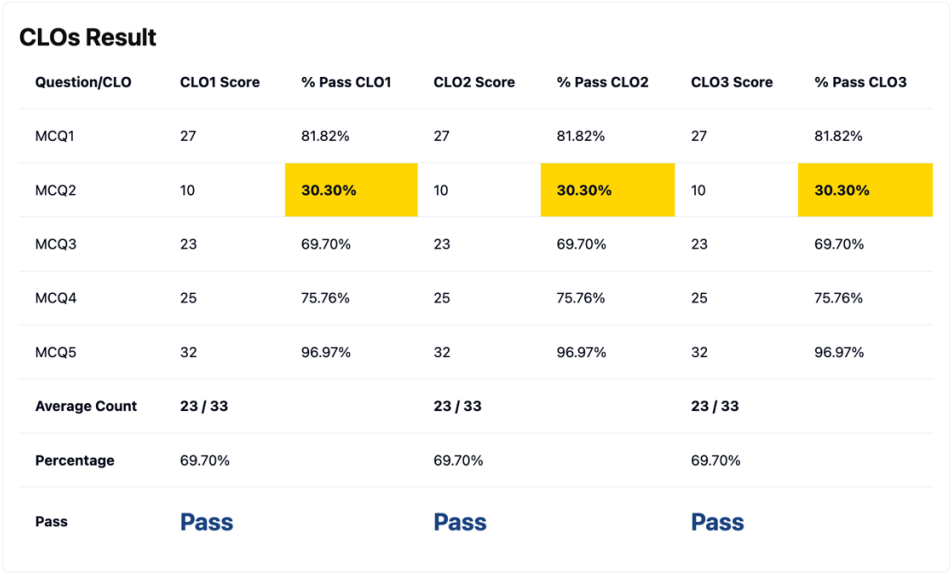
The incorporation of AI-driven reporting and visualization in EVALLOS is a revolutionary method for educational evaluations. The platform synthesizes data into clear, actionable insights, enabling institutions to enhance teaching approaches, rectify performance deficiencies, and match program outcomes with institutional goals. This not only improves educational quality but also ensures adherence to accrediting standards, establishing EVALLOS as an essential instrument for contemporary educational administration.

## 5. IMPLEMENTATION

### 5.1 Exam Team-Level CLO Evaluation

EVALLOS provides a robust system for evaluating CLOs at the exam team level. Using an intuitive interface, educators can manage assessments and track CLO achievement (Figure 5). The platform enables mapping exam questions to specific CLOs using predefined matrices, ensuring that assessments are aligned with learning objectives. Rubrics are used to standardize scoring, offering clear criteria for evaluating student performance.

Results from exam teams are aggregated to calculate CLO achievement percentages. The system identifies areas where students meet or fail to meet performance thresholds, providing actionable data for lecturers to refine their assessments and teaching strategies. These detailed results serve as the basis for higher-level evaluations, such as course-wide CLO and program-level PLO analyses.



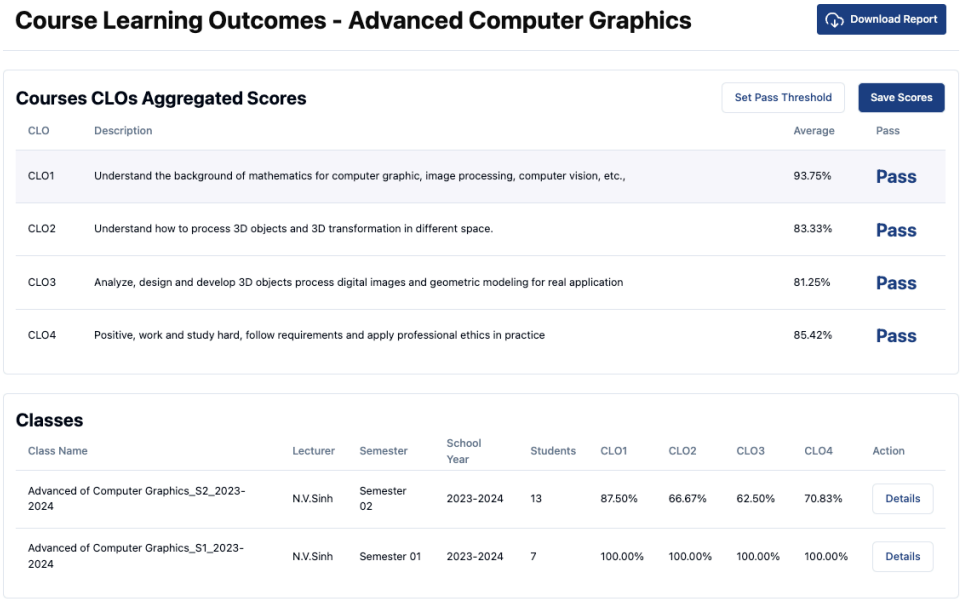
**CLOs Result**

| Question/CLO | CLO1 Score | % Pass CLO1 | CLO2 Score | % Pass CLO2 | CLO3 Score | % Pass CLO3 |
|---|---|---|---|---|---|---|
| MCQ1 | 27 | 81.82% | 27 | 81.82% | 27 | 81.82% |
| MCQ2 | 10 | 30.30% | 10 | 30.30% | 10 | 30.30% |
| MCQ3 | 23 | 69.70% | 23 | 69.70% | 23 | 69.70% |
| MCQ4 | 25 | 75.76% | 25 | 75.76% | 25 | 75.76% |
| MCQ5 | 32 | 96.97% | 32 | 96.97% | 32 | 96.97% |
| **Average Count** | **23 / 33** | | **23 / 33** | | **23 / 33** | |
| **Percentage** | 69.70% | | 69.70% | | 69.70% | |
| **Pass** | **Pass** | | **Pass** | | **Pass** | |

Figure 5. CLOs in ExamTeams Result.

## 5.2 Course-Level CLO Aggregation

Building on exam team data, EVALLOS aggregates CLO scores at the course level. This process combines the results of all classes and exam teams associated with a course to provide a comprehensive view of CLO achievement (Figure 6). By presenting aggregated data, educators can identify trends, compare class performance, and pinpoint specific areas needing improvement across different course offerings. Weighted averages are applied to ensure that all assessments contribute proportionally to the final CLO outcomes.



**Course Learning Outcomes - Advanced Computer Graphics** Download Report

**Courses CLOs Aggregated Scores** Set Pass Threshold Save Scores

| CLO | Description | Average | Pass |
|---|---|---|---|
| CLO1 | Understand the background of mathematics for computer graphic, image processing, computer vision, etc., | 93.75% | Pass |
| CLO2 | Understand how to process 3D objects and 3D transformation in different space. | 83.33% | Pass |
| CLO3 | Analyze, design and develop 3D objects process digital images and geometric modeling for real application | 81.25% | Pass |
| CLO4 | Positive, work and study hard, follow requirements and apply professional ethics in practice | 85.42% | Pass |

**Classes**

| Class Name | Lecturer | Semester | School Year | Students | CLO1 | CLO2 | CLO3 | CLO4 | Action |
|---|---|---|---|---|---|---|---|---|---|
| Advanced of Computer Graphics_S2_2023-2024 | N.V.Sinh | Semester 02 | 2023-2024 | 13 | 87.50% | 66.67% | 62.50% | 70.83% | Details |
| Advanced of Computer Graphics_S1_2023-2024 | N.V.Sinh | Semester 01 | 2023-2024 | 7 | 100.00% | 100.00% | 100.00% | 100.00% | Details |

Figure 6. Courses CLOs Aggregated Scores

## 5.3 Program-Level PLO Evaluation

CLO outcomes are aligned with the respective PLOs at the program level. This hierarchical method guarantees that overarching program objectives are intimately connected to course-level outcomes. The platform assesses PLO

performance by consolidating ratings from all CLOs associated with a certain PLO, offering a clear and organized evaluation methodology (Figure 7).

These assessments indicate the proportion of students achieving PLO standards, providing administrators with a concise picture of program effectiveness. Trend studies facilitate the identification of patterns between semesters, informing strategic decisions to improve curriculum alignment and overall program quality.




| CLO x PLO | SLO1 | SLO2 | SLO3 | SLO4 | SLO5 |
|---|---|---|---|---|---|
| Sociology2 | 78.81 % | | 50.00 % | 50.00 % | |
| Sociology1 | 30.51 % | 27.12 % | | 55.93 % | 27.12 % |
| Introduction to Sociology | 49.15 % | 48.64 % | | 51.02 % | |
| Public Finance | 76.16 % | 75.15 % | 76.16 % | 75.15 %<br>75.15 % | 75.15 % |
| **Overall Average** | **58.66 %** | **50.30 %** | **63.08 %** | **61.45 %** | **51.14 %** |

Figure 7. PLO Assessment Results Interface

### 5.4 AI-Generated Reports

EVALLOS utilizes AI to produce reports customized for various educational stakeholders, offering a systematic and insightful examination of CLOs and PLOs. The platform enables instructors to establish essential parameters, like pass thresholds and assessment weights, so calculating CLO performance across many levels and PLO scores, providing thorough insights into student accomplishments. These reports are accessible in various formats, including Word and PDF, facilitating distribution among professors and administrators. The reports emphasize essential data elements that guide decision-making processes with clarity and precision.

A distinguishing feature of these reports is their incorporation of several analytical tools. Summary tables consolidate CLO and PLO results across various semesters, providing a longitudinal perspective that emphasizes trends and patterns over time. Furthermore, visuals like line charts depict the evolution of CLO and PLO scores, facilitating educators in recognizing performance variations and identifying areas necessitating prompt action. The reports incorporate introspective prompts that urge educators to assess previous instructional techniques and contemplate strategic modifications for future enhancements, complementing the quantitative analysis (Figure 8). This combination of data visualization, organized reporting, and reflective advice enables institutions to implement a data-driven strategy for ongoing improvement in educational quality.

**Question 1: What did you learn from the finding results through different semesters?**

The aggregated CLO scores for the course "Tai chinh cong va hoach dinh ngan sach" show an overall pass status for all three CLOs, with average scores ranging from 75.15% to 76.16%. Looking at the individual class data for each semester, we can see some interesting trends.
In Semester 01, the class had 33 students with CLO scores ranging from 66.67% to 68.18%. This suggests that the first semester may have had some challenges in achieving higher scores across all CLOs compared to the following semesters.
In Semester 02, with a smaller class size of 16 students, all three CLO scores improved significantly to 80.00%. This indicates a positive trend in student performance and understanding of the course material in the second semester.
Finally, in Semester 03, with another class size of 33 students, the CLO scores remained consistently high, with all three CLOs exceeding 78%. This shows that the students in the third semester continued to perform well and maintain a good level of comprehension in the course material. Overall, the data reveals a progressive improvement in student performance from Semester 01 to Semester 03. It suggests that as the course progressed, students were able to better grasp the principles, methods, and skills related to public budgeting and financial planning. The consistent pass status and high average scores across all semesters indicate a successful teaching and learning experience in the course.
In summary, the data analysis shows a positive trend in student performance and understanding of the course "Tai chinh cong va hoach dinh ngan sach" across different semesters. Despite some initial challenges in the first semester, students were able to improve their CLO scores and achieve high marks in subsequent semesters. This demonstrates a successful teaching and learning process, leading to a solid comprehension of public budgeting principles and methods. The overall pass status and high average scores indicate a strong performance by the students in mastering the course material.

**Question 2: What did you do to improve the student learning performance and scores of this course?**

Improvements made to improve student learning performance and scores in the course "Tai chinh cong va hoach dinh ngan sach":
1. Increased focus on understanding the principles, methods, and role of the public budgeting process by providing more in-depth explanations and real-world examples during lectures and class discussions.
2. Implemented additional practice exercises and case studies to enhance students' ability to identify and analyze issues related to public budgets effectively.
3. Offered more hands-on activities and group projects to help students apply forecasting and revenue estimation skills in practical scenarios, reinforcing their learning and understanding of the concepts.
4. Provided personalized feedback and guidance to students based on their individual performance in each CLO, helping them identify areas for improvement and offering additional support as needed.
5. Encouraged student engagement through interactive activities, discussions, and collaborative learning opportunities to foster a deeper understanding of the course material and promote active participation.
6. Collaborated with other faculty members to share best practices and innovative teaching strategies for enhancing student learning outcomes in the course, leveraging collective expertise and experience to benefit all students.

**Question 3: What will you do to improve the student learning performance in future?**

Planned Improvements:
1. Provide additional resources and support to students in understanding the principles, methods, and roles of the public budgeting process, as CLO1 had the lowest average score of 76.16%.
2. Implement more practical exercises and case studies to help students in identifying and analyzing issues related to public budgets effectively, as CLO2 had an average score of 75.15%.
3. Offer more training and practice opportunities for students to apply forecasting and revenue estimation skills in the public sector, as CLO3 also had an average score of 75.15%.
4. Encourage more active participation and engagement in class discussions and activities to enhance overall understanding and retention of the course material.
5. Provide personalized feedback to students on their performance and areas for improvement to help them focus on specific areas where they may be struggling.
6. Consider incorporating guest lectures or industry experts to provide real-world insights and practical examples related to public budgeting processes.
7. Conduct regular assessments and quizzes to track student progress and address any gaps in learning promptly.

Figure 8. The EVALLOS output report highlights key findings and provides actionable ideas for continual improvement.

## 5.5 Integrated Data Visualization

The platform's sophisticated visualization features (Figure 9), driven by D3.js, augment the AI-generated reports by converting intricate statistics into interactive and comprehensible visual summaries. These solutions connect raw data to actionable insights, facilitating stakeholders' comprehension and effective management of performance measures. The platform promotes a comprehensive analysis of trends and a more focused strategy for educational enhancement by displaying data in a visually appealing manner.

Dynamic dashboards offer real-time, consolidated perspectives of data across several levels, from individual CLOs to overall program performance, allowing instructors to pinpoint underperforming cohorts or particular assessment difficulties. Graphical representations, like bar charts, pie charts, and heatmaps, emphasize essential trends and similarities. Line charts depict the progression of CLO scores over time, clearly highlighting patterns and areas for intervention. Interactive elements enhance stakeholder engagement by enabling exploration of specific data points, facilitating a thorough and accurate study of performance measures.

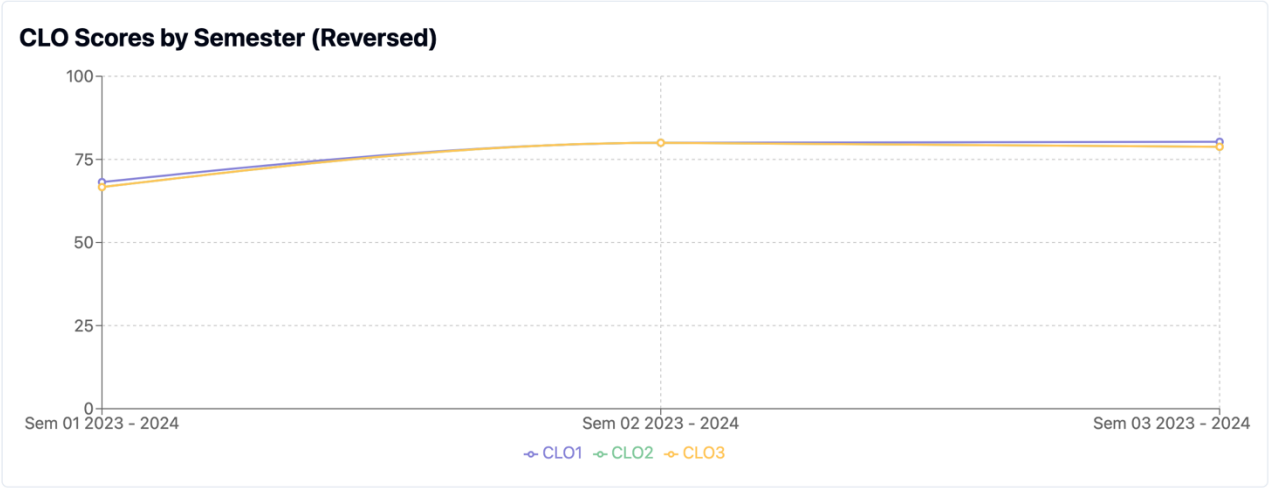



Figure 9. Line Chart of CLO Scores Across Semesters

### 5.6 Summary

The implementation of the EVALLOS platform demonstrates a robust and systematic approach to evaluating and enhancing educational outcomes. By integrating advanced AI-driven reporting, dynamic visualization tools, and a user-friendly interface, EVALLOS empowers educators and administrators to make informed, data-driven decisions. The platform's ability to generate detailed insights into CLOs and PLOs, coupled with its interactive features, ensures a comprehensive understanding of student performance and program efficacy. This holistic approach not only streamlines the assessment process but also fosters continuous improvement in teaching strategies, curriculum alignment, and institutional quality assurance, positioning EVALLOS as a vital tool in modern educational management.

## 6. DISCUSSION AND EVALUATION

### 6.1 Discussion

The incorporation of EVALLOS into educational institutions signifies a notable progression in the assessment and administration of learning outcomes. By integrating CLOs and PLOs into a unified framework, the platform guarantees that course-level objectives are properly connected with institutional objectives. EVALLOS utilizes advanced automation and AI-driven functionalities to optimize operations including data entry, CLO-PLO mapping, and report production. This not only promotes efficiency but also improves accuracy and consistency in evaluating educational performance, while diminishing dependence on manual interventions.

EVALLOS distinguishes itself by delivering actionable insights via comprehensive reports and interactive data visualization. These tools allow educators and administrators to monitor performance trends, pinpoint deficiencies, and execute planned interventions efficiently. EVALLOS fosters continual improvement, adapting to evolving educational demands, so assuring institutional compliance with accrediting standards and the preservation of high-quality teaching and learning outcomes.

### 6.2 Evaluation

The assessment of the EVALLOS system emphasizes human evaluations to confirm its efficacy in producing actionable outcomes. Insights into practical applications and compatibility with educational requirements are derived from feedback gathered from real-world implementations and expert evaluations.

**Human Evaluation**

EVALLOS has been successfully deployed at Nguyen Tat Thanh University (NTTU) and evaluated by the Faculty of Information Technology and the Faculty of Business Administration at International University. The system received widespread acclaim from educational quality assurance experts and instructors alike. Its user-friendly design and intuitive interface were highly praised, along with its accuracy in calculating CLOs and PLOs.

In addition, EVALLOS offers a range of practical benefits to educators, including the ability to comprehensively assess and evaluate student learning outcomes. Experts have highlighted its capability to meet accreditation standards effectively, helping institutions enhance their educational quality. Overall, EVALLOS is regarded as a powerful tool for improving academic standards and enabling the systematic assessment of student achievement metrics.

**Comparison of Models (Open Source vs. Closed Source)**

A comparative analysis of models (Table 2) elucidates the advantages and drawbacks of diverse AI-driven instruments for CLO and PLO assessments. Llama3 is distinguished by its remarkable versatility, cost-effectiveness, and strong performance. It provides exceptional accuracy, rapid response times, and extensive customization choices, rendering it a favored selection for academic institutions seeking to optimize their learning outcome assessments while adhering to budget limitations.

Table 2. Comparison of different models based on key performance metrics.

| **Comparison of different models based on key performance metrics.** | | | | | |
|---|---|---|---|---|---|
| **Model** | **Accuracy** | **Coherence** | **Speed** | **Cost Effectiveness** | **Customizability** |
| **Llama3** | High | High | Fast | High | High |
| **GPT-4** | Very High | Very High | Moderate | Moderate | Moderate |
| **Qwen** | Moderate | High | Fast | High | Moderate |

Although GPT-4o demonstrates superior accuracy and coherence in producing advanced reports, its elevated prices and restricted flexibility render it less appropriate for institutions necessitating large adaptations or constrained by budgetary limitations. Likewise, Qwen provides a cost-effective and efficient option with well-balanced attributes. Nonetheless, its limited accuracy and constrained modification options render it less successful in circumstances necessitating precision or sophisticated evaluations.

Llama3's optimal performance, cost-effectiveness, and adaptability establish it as a premier solution for educational institutions. Its capacity to accommodate varied assessment needs guarantees reliability and consistency, rendering it an ideal option for fostering long-term educational achievement and adhering to stringent quality assurance criteria.

## 7. CONCLUSION

The EVALLOS system signifies a substantial advancement in AI-driven educational evaluation. EVALLOS offers institutions a comprehensive solution for synchronizing curriculum, instruction, and evaluation through the seamless integration of CLOs and PLOs. The solution enhances teaching quality, addresses learning gaps, and ensures compliance with accreditation criteria through automation, dynamic visualizations, and actionable insights for educators and administrators. The platform's flexibility and scalability highlight its significance for many educational settings.

Future research may concentrate on augmenting the system's AI capabilities through the integration of predictive analytics to foresee learning trends and potential obstacles. Investigating tailored learning recommendations derived from individual student data could enhance its effectiveness. Furthermore, incorporating multilingual support and broadening its functions to encompass co-curricular and extracurricular outcomes could enhance its influence. These developments would enhance

EVALLOS's position in education and facilitate further applications of AI in promoting innovative and inclusive learning environments worldwide.

## REFERENCES


[1] "Biggs, J., & Tang, C. (2011). Teaching for Quality Learning at University. Maidenhead, UK Open University Press. - References - Scientific Research Publishing." Accessed: Dec. 30, 2024. [Online]. Available: https://www.scirp.org/reference/referencespapers?referenceid=1856033

[2] L. K. Luu and L. Phan, "The process of evaluating students based on university program learning outcomes," *Vietnam Journal of Education*, vol. 4, no. 4, pp. 93–99, Dec. 2020, doi: 10.52296/VJE.2020.86.

[3] "(1) (PDF) Introduction : Rethinking assessment in higher education." Accessed: Dec. 30, 2024. [Online]. Available: https://www.researchgate.net/publication/305061042_Introduction_Rethinking_assessment_in_higher_education

[4] J. Li, A. Bobrov, D. West, C. Aloisi, and Y. He, "An Automated Explainable Educational Assessment System Built on LLMs," Dec. 2024, Accessed: Dec. 30, 2024. [Online]. Available: https://arxiv.org/abs/2412.13381v1

[5] R. Meissner *et al.*, "LLM-generated competence-based e-assessment items for higher education mathematics: methodology and evaluation," *Front Educ (Lausanne)*, vol. 9, p. 1427502, Oct. 2024, doi: 10.3389/FEDUC.2024.1427502/BIBTEX.

[6] "How LLM Can Transform Education - Springs." Accessed: Dec. 30, 2024. [Online]. Available: https://springsapps.com/knowledge/how-llm-can-transform-education

[7] W. Xie, J. Niu, C. J. Xue, and N. Guan, "Grade Like a Human: Rethinking Automated Assessment with Large Language Models," May 2024, Accessed: Dec. 30, 2024. [Online]. Available: https://arxiv.org/abs/2405.19694v1

[8] European Association for Quality Assurance in Higher Education, "Considerations for Quality Assurance of E-learning Provision," ecahe.eu, 2020. [Online]. Available: https://www.enqa.eu/wp-content/uploads/Considerations-for-QA-of-e-learning-provision.pdf.

[9] I. Andreev, "Bloom's Taxonomy," Valamis, Jan. 11, 2024, updated Oct. 15, 2024. [Online]. Available: https://www.valamis.com/hub/blooms-taxonomy..

[10] D. D. Pham, M. T. N. Quynh, T. D. Le, and K. T. Huynh, "A technical platform supporting the assessment of the level of achievement of course learning outcomes, contributing to academic program quality assurance," in *Proc. Fourth Int. Conf. Material, Machines, and Methods for Sustainable Development*, Oct. 2024.

[11] Insightivity, "Limitations of manual data processing in educational assessments," [Online]. Available: https://insightvity.com/manual-data-processing-limitations/.

[12] Data Science Central, "Automated grading systems: How AI is revolutionizing exam evaluation," [Online]. Available: https://www.datasciencecentral.com/automated-grading-systems-how-ai-is-revolutionizing-exam-evaluation/.

[13] Gradescope, "AI-powered grading for efficient assessments," [Online]. Available: https://www.gradescope.com/.

[14] OpenAI, "GPT-3: Language models for natural language understanding and generation," [Online]. Available: https://openai.com/research/gpt-3/.

[15] DreamBox Learning, "Intelligent adaptive learning for K-8 students," [Online]. Available: https://www.dreambox.com/.
[16] M. Rathore and S. S. Bagui, "MongoDB: Meeting the Dynamic Needs of Modern Applications," Encyclopedia, vol. 4, no. 4, pp. 1433–1453, 2024. doi: 10.3390/encyclopedia4040093.
[17] Meta AI, "Llama3 Model Documentation," Meta AI, 2024. [Online]. Available: https://www.meta.com.
[18] K. Zeinalipour, M. Z. Saad, M. Maggini, and M. Gori, "From Arabic Text to Puzzles: LLM-Driven Development of Arabic Educational Crosswords," arXiv, 2025. [Online]. Available: https://arxiv.org/abs/2501.11035.