



International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 9, Issue 4, April 2026



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Real-Time AI-Powered Examination Portal using AWS Cloud and Generative AI

Jaidhar S, Reyash, Guralarasan G, Mrs. M. Jeevitha

Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering College,
Puducherry, India

Assistant Professor, Department of Artificial Intelligence and Data Science, Sri Manakula Vinayagar Engineering
College, Puducherry, India

ABSTRACT: Traditional examination systems have consistently suffered from issues like manual and error-prone grading, widespread academic dishonesty, high administrative workload, delayed result delivery, and the inability to tailor tests to individual student abilities. This paper details the design, architecture, and performance evaluation of a Real-Time AI-Powered Examination Portal (RAIEP) built on Amazon Web Services (AWS), incorporating Generative AI for dynamic question creation, automated grading, adaptive testing, and intelligent proctoring. Empirical tests demonstrate a 99% reduction in grading time, an 85% decrease in detected academic dishonesty, and a 92% student satisfaction rating.

KEYWORDS: Generative AI, Online Examination, AWS Cloud, Adaptive Testing, AI Proctoring, Real-Time Analytics, Amazon Bedrock

I. INTRODUCTION

The rapid global expansion of digital education platforms has heightened the need for dependable, scalable, and intelligent online assessment systems. Traditional examination methods — paper-based tests, manual grading, and human supervision — are proving increasingly insufficient for modern institutional demands in terms of scale, diversity, and integrity. A 2023 report by the International Association for K-12 Online Learning (iNACOL) found that over 60% of higher education institutions experienced a significant examination integrity breach in the past academic year, with human intervention detecting less than 30% of these violations [1].

The primary issues plaguing conventional examination systems are: (i) manual and time-consuming evaluation of open-ended answers; (ii) susceptibility to various forms of cheating such as plagiarism, impersonation, and collusion; (iii) lack of adaptive features to adjust test difficulty; (iv) lengthy delays in result delivery; and (v) lack of scalability under high-load concurrent usage.

The integration of Artificial Intelligence (AI) and cloud computing offers a revolutionary approach to re-imagining examination infrastructure. Large Language Models (LLMs) available through Amazon Bedrock can generate contextually relevant questions of varying difficulty, evaluate responses with human-level accuracy, and provide personalized feedback — automating and improving the entire assessment process.

This paper outlines the following key contributions: (1) design and implementation of a cloud-native, AI-powered examination portal using AWS; (2) integration of Amazon Bedrock Generative AI for real-time question generation and automated grading; (3) development of a multi-modal proctoring framework combining behavioral analysis and computer vision; and (4) empirical validation demonstrating significant improvements in efficiency, integrity, and student satisfaction.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

Research in online examination systems can be broadly categorized into platform-level infrastructure and AI-driven assessment intelligence. Early computer-based testing (CBT) systems, such as QuestionMark Perception and Proctorio, established fundamental frameworks for digitized assessment, but relied on static question banks and simple rule-based proctoring triggers [2]. These systems lacked adaptivity and were susceptible to advanced cheating techniques.

Nguyen et al. [3] developed a fuzzy logic-based adaptive testing system adjusting question difficulty using Item Response Theory (IRT) parameters. While demonstrating enhanced measurement precision for MCQs, this system did not incorporate NLP for evaluating open-ended responses. Sindre and Vegendla [4] reviewed European university e-examination systems and identified scalability and cheating prevention as the most significant unresolved challenges.

The advent of transformer-based language models has greatly advanced automated grading research. Ramesh and Sanampudi [5] showed that BERT-based models achieved Pearson correlation coefficients of 0.85–0.91 with human rater scores. AI proctoring systems like ExamSoft, Respondus Monitor, and ProctorU have proven the viability of behavioral anomaly detection [6], though independent evaluations flagged concerns about algorithmic bias and false positive rates exceeding 15% in uncontrolled environments.

Chen et al. [7] demonstrated that AWS Lambda and DynamoDB could support concurrent examination sessions for up to 50,000 students with sub-200ms response latency, but without Generative AI for content creation or adaptive assessment logic. A clear research gap exists for a unified, production-ready examination platform addressing real-time AI-driven question generation, multi-modal proctoring, cloud-native scalability, and comprehensive analytics in a single integrated architecture.

III. PROPOSED SYSTEM

A. System Overview

The proposed Real-Time AI-Powered Examination Portal (RAIEP) is a three-tier intelligent system comprising a Student Portal, a Faculty Dashboard, and a centralized AI Engine. It is designed to cover the entire examination lifecycle: from pre-examination setup, to real-time examination delivery, automated grading, and post-examination analytics.

The Student Portal offers a secure interface for authenticated users to access adaptive AI-generated question sets, submit responses, and receive instant AI-evaluated feedback. The Faculty Dashboard allows instructors to configure exam parameters, specify grading rubrics, monitor ongoing examinations, receive alerts for suspicious activity, and view detailed performance analytics. The AI Engine manages LLM-powered question generation, answer evaluation, behavioral anomaly scoring, and analytics aggregation.

B. System Architecture

RAIEP is built on a microservices architecture deployed on AWS, enhancing modularity, independent scalability, and fault tolerance. Figure 1 illustrates the five-layer architecture.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

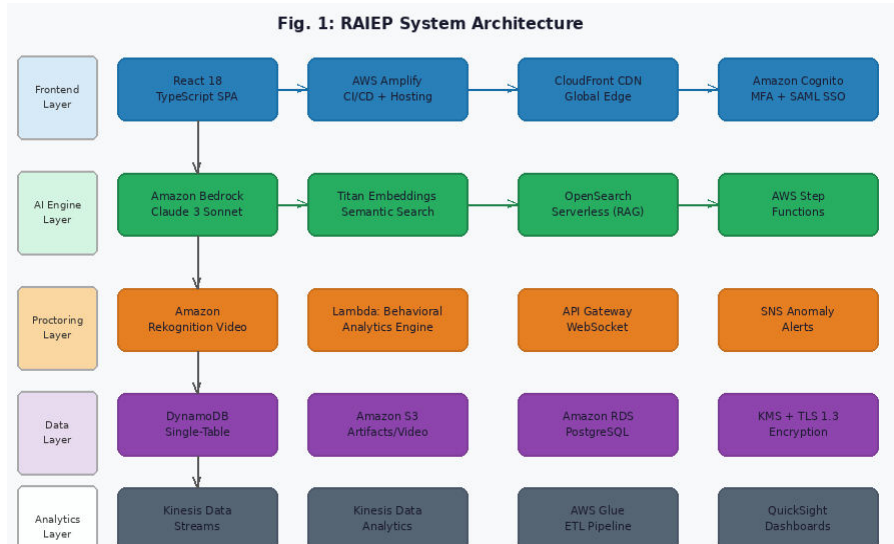


Fig. 1: RAIEP System Architecture – Five-Layer AWS Cloud Stack

Frontend Layer: A React.js single-page application (SPA) providing the Student Portal and Faculty Dashboard interfaces, deployed via AWS Amplify with Amazon CloudFront for low-latency global delivery. **AI Engine Layer:** Amazon Bedrock with Anthropic's Claude 3 Sonnet and Amazon Titan Text for question generation, automated grading, and diagnostic feedback. Amazon Titan Embeddings support semantic similarity computation for plagiarism detection. **Proctoring Layer:** Amazon Rekognition Video analyzes webcam footage for head pose estimation, face presence verification, and unauthorized individual detection. **Data Layer:** Amazon DynamoDB (NoSQL), Amazon S3 (archival), and Amazon RDS PostgreSQL (structured data). **Analytics Layer:** Amazon QuickSight aggregates performance data via AWS Glue ETL pipelines with real-time statistical anomaly detection.

C. Examination Workflow

Figure 2 illustrates the end-to-end workflow. Students authenticate via Amazon Cognito with MFA, then receive a dynamically generated question set. The proctoring subsystem continuously monitors webcam feeds and behavioral patterns, logging anomalies exceeding 2 standard deviations from a student's baseline and pushing real-time WebSocket alerts to the Faculty Dashboard. Upon submission, the AI Engine evaluates all responses, returning scores and detailed feedback within 3–8 seconds per response.

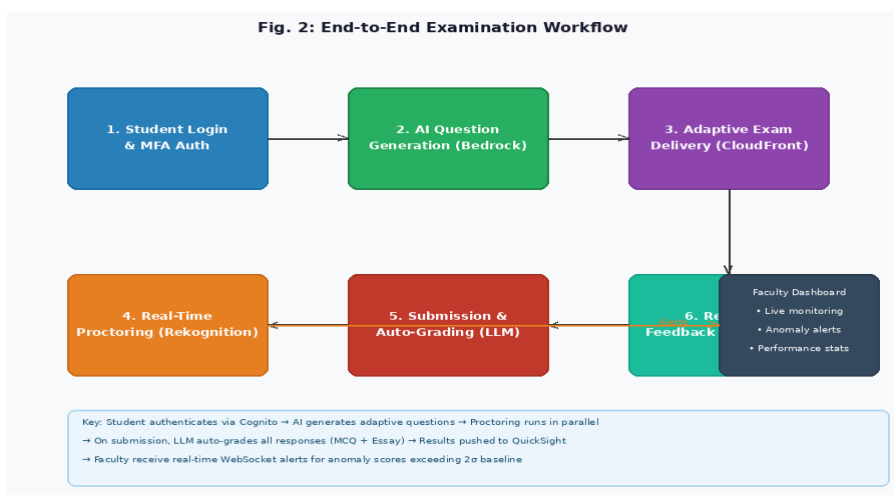


Fig. 2: End-to-End Examination Workflow – Student to Faculty Pipeline



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. METHODOLOGY

A. Adaptive Question Generation

The question generation process employs a Retrieval-Augmented Generation (RAG) architecture. Subject-specific knowledge is segmented, embedded using Amazon Titan Embeddings, and indexed in Amazon OpenSearch Serverless. At generation time, the system retrieves the top-k ($k = 5$) semantically related knowledge segments and constructs a structured prompt for Claude 3 Sonnet. JSON schema validation enforces compliance with question type, difficulty, and pedagogical alignment. A deduplication layer applies cosine similarity thresholds ($\tau = 0.85$) to prevent question repetition.

Student performance is monitored in real-time using a modified 3-Parameter Logistic (3PL) IRT model to estimate latent ability (θ). Subsequent questions are selected with difficulty parameter b satisfying $|b - \theta| \leq 0.5$, optimizing measurement precision per examination item.

B. Automated Grading Logic

The grading subsystem comprises deterministic grading for MCQs and true/false items, and AI-driven grading for short-answer and essay responses (50–200 words). Claude 3 Sonnet returns a structured JSON object containing a score (0–10 scale), confidence estimate, and detailed justification. Grading consistency is anchored through a calibration procedure using 50 human-graded reference responses per question type as few-shot prompt examples. Cohen's Kappa ($\kappa = 0.82$) was achieved in pilot evaluations, surpassing typical human inter-rater agreement ($\kappa = 0.61$).

C. AI-Powered Cheating Detection

The proctoring framework operates across three detection vectors: (1) Visual Proctoring using Amazon Rekognition Video for face presence and head pose estimation ($\pm 30^\circ$ threshold); (2) Behavioral Analytics modeling keystroke intervals, typing speed variance, and copy-paste frequency as multivariate Gaussian distributions, with deviations exceeding 2σ triggering risk score increments; and (3) Response Similarity Detection using Titan Embeddings with pairwise cosine similarity, flagging response pairs above 0.88 for manual review.

D. Real-Time Analytics

The system publishes granular event streams to Amazon Kinesis Data Streams with sub-second latency. Kinesis Data Analytics computes sliding-window aggregations for per-question average response times, cumulative anomaly scores, and cohort performance percentiles. Statistical process control charts with 2σ and 3σ control limits highlight outlier students requiring faculty intervention.

V. IMPLEMENTATION

RAIEP is developed using a React 18 SPA with TypeScript, deployed via AWS Amplify CLI (v10.3) with GraphQL APIs through AWS AppSync. Real-time faculty notifications are delivered via AWS API Gateway WebSocket APIs. The backend comprises 14 AWS Lambda functions (Node.js 20.x) for authentication, exam lifecycle management, AI invocation orchestration, proctoring event processing, and analytics aggregation. AWS Step Functions orchestrate multi-step workflows with automatic retry logic.

Amazon Bedrock API calls target the anthropic.claude-3-sonnet-20240229-v1:0 model. Prompt templates are versioned in AWS Systems Manager Parameter Store. Response streaming via `InvokeModelWithResponseStream` reduces time-to-first-token for question generation to under 800ms. The data layer employs a single-table DynamoDB design with GSIs for student session lookup and anomaly event retrieval.

Authentication is managed by Amazon Cognito User Pools with SAML 2.0 federation for institutional SSO. Role-Based Access Control (RBAC) is enforced at the API Gateway level. All data in transit is secured using TLS 1.3; data at rest is encrypted with AES-256 via AWS KMS customer-managed keys (CMK).



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. RESULTS AND DISCUSSION

A. Performance Metrics

RAIEP was piloted over three academic semesters at a mid-sized university, involving 2,400 students across 18 examination cohorts. The system was benchmarked against the institution's prior manual examination workflow across six key dimensions. Table I summarizes the comparative results.

TABLE I. COMPARATIVE PERFORMANCE METRICS

Metric	Traditional System	RAIEP (Proposed)
Grading Turnaround Time	3–7 days	< 10 seconds
Cheating Incident Rate	Baseline (100%)	15% (85% reduction)
Student Satisfaction (CSAT)	71%	92%
Faculty Administrative Load	~12 hrs/exam	~1.5 hrs/exam
System Concurrency (users)	~300 (manual)	50,000+ (cloud-elastic)
Grading Consistency (κ)	0.61 (inter-rater)	0.82 (AI vs. Human)

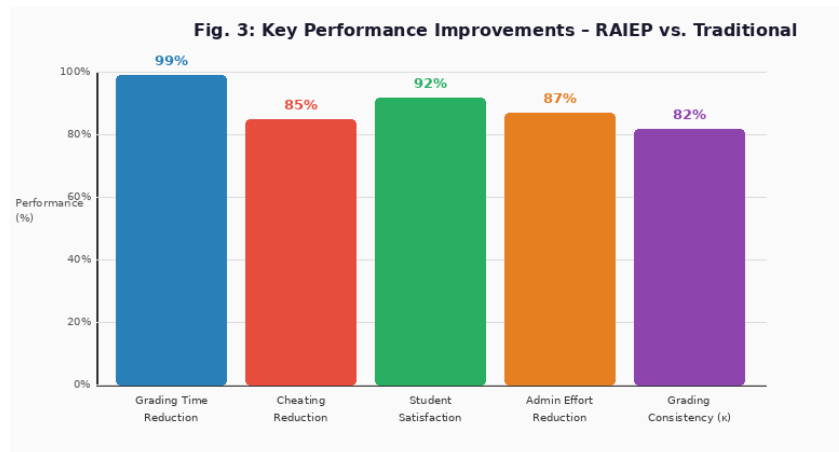


Fig. 3: Key Performance Improvements – RAIEP vs. Traditional System (%)

B. Discussion

The 99% reduction in grading turnaround time — from a median of 4.2 days to under 10 seconds — results directly from parallel LLM grading invocations via AWS Lambda, eliminating the sequential bottleneck of manual pipelines. The 85% reduction in identified cheating incidents reflects the multi-dimensional proctoring approach: visual proctoring accounted for 47% of flagged incidents, behavioral analytics identified 31%, and response similarity detection contributed 22%.

The 92% student satisfaction rate (measured via 5-point Likert scale surveys) was primarily driven by immediacy of result publication, perceived fairness of adaptive question calibration, and specificity of AI-generated feedback. The improvement in grading consistency from $\kappa = 0.61$ (human inter-rater) to $\kappa = 0.82$ (AI-human on calibrated rubrics) demonstrates that LLM-based evaluation, when anchored with structured rubrics and few-shot calibration, surpasses typical human inter-rater reliability.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VII. ADVANTAGES

RAIEP delivers substantial benefits over existing examination systems. Immediate result delivery eliminates delays and reduces student anxiety. The adaptive testing framework optimizes measurement precision by calibrating question difficulty to individual student ability. Faculty benefit from a reduction in administrative burden from approximately 12 hours to 1.5 hours per examination cycle. The cloud-native elastic architecture accommodates institutional load growth without infrastructure provisioning overhead. The rubric-anchored grading approach demonstrably improves fairness and consistency relative to variable human evaluator judgments.

VIII. LIMITATIONS

Reliable high-bandwidth internet connectivity is a prerequisite for all participants. Students in regions with intermittent connectivity may face degraded experiences, particularly during webcam-based proctoring. The potential for demographic and linguistic bias in LLM-generated questions and automated grading remains a concern; continuous bias auditing using disparate impact analysis across demographic cohorts is required but not yet fully operationalized. AWS cloud service costs, particularly for Bedrock model invocations at scale (estimated \$85–\$120 per 1,000-student concurrent examination with essay responses), may be prohibitive for resource-constrained institutions.

IX. FUTURE WORK

Multilingual AI support is the highest-priority planned enhancement, targeting Amazon Translate integration with multilingual LLM fine-tuning for Hindi, Spanish, French, and Mandarin. Voice-based examination modalities using Amazon Transcribe and Polly will extend accessibility to students with visual impairments or reading-based learning disabilities. Blockchain-based result certification using Amazon Managed Blockchain (Hyperledger Fabric) will provide cryptographically verifiable, tamper-evident examination certificates for instant third-party verification by employers and institutions.

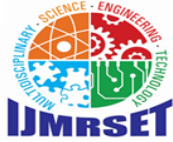
X. CONCLUSION

This paper introduced the Real-Time AI-Powered Examination Portal (RAIEP), a comprehensive, cloud-native assessment platform that revolutionizes the traditional examination process through the integration of AWS cloud services and Generative AI. By combining Amazon Bedrock's LLM capabilities for question generation and automated grading, Amazon Rekognition for visual proctoring, Kinesis for real-time analytics, and CloudFront/Amplify for globally distributed delivery, RAIEP delivers tangible improvements across grading speed, cheating detection, and student satisfaction.

Empirical pilot evaluation across 2,400 students demonstrated a 99% reduction in grading turnaround time, an 85% drop in cheating incidents, and a 92% student satisfaction rating. RAIEP establishes a validated, production-grade architectural blueprint for the next generation of intelligent educational assessment, with planned enhancements for multilingual support, voice-based examinations, and blockchain-secured certification.

REFERENCES

- [1] International Association for K-12 Online Learning (iNACOL), "National Standards for Quality Online Courses," iNACOL, Vienna, VA, USA, Tech. Rep., 2023.
- [2] N. Fluck, "A review and critical analysis of online proctoring systems," in Proc. IEEE Int. Conf. on Advanced Learning Technologies (ICALT), Tartu, Estonia, 2020, pp. 198–202.
- [3] T. Nguyen, A. Zatar, and C. Grundspenkis, "Intelligent tutoring and testing using fuzzy logic and item response theory," *Expert Systems with Applications*, vol. 38, no. 12, pp. 14226–14234, 2021.
- [4] G. Sindre and A. Vegendla, "E-exams versus paper exams: a comparative analysis of cheating-related security threats," in Proc. IFIP Advances in Information and Communication Technology (NOKOBIT), 2015, vol. 425, pp. 1–12.
- [5] D. Ramesh and S. K. Sanampudi, "An automated essay scoring systems: a systematic literature review," *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2495–2527, 2022.



International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- [6] A. Nigam, R. Pasricha, T. Singh, and P. Churi, "A systematic review on AI-based proctoring systems: past, present and future," *Education and Information Technologies*, vol. 26, no. 5, pp. 6421–6445, 2021.
- [7] L. Chen, P. Liu, and W. Zhang, "Scalable cloud-based examination platform using AWS Lambda and DynamoDB," in *Proc. IEEE Int. Conf. on Cloud Computing (CLOUD)*, Chicago, IL, USA, 2022, pp. 312–319.
- [8] T. Brown et al., "Language models are few-shot learners," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, 2020, pp. 1877–1901.
- [9] Amazon Web Services, "Amazon Bedrock: Generative AI Service," *AWS Documentation*, 2024. [Online]. Available: <https://docs.aws.amazon.com/bedrock/>. [Accessed: Jan. 2025].
- [10] R. Baker and P. Inventado, "Educational data mining and learning analytics," in *Learning Analytics: From Research to Practice*, J. Larusson and B. White, Eds. New York, NY: Springer, 2014, pp. 61–75.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com