Published by The College of Science & Information Technology (COSIT), TASFUED, Vol. 20, No. 1, pp. 15-26.

COSIT, TASFUED
Journal of Science and Information Technology (JOSIT)

FedExamGuard: A Privacy-Preserving AI Framework for Intelligent Examination Governance

Arowolo, O. A.
Department of Computer Science, Tai Solarin Federal University of Education, Ijagun, Ogun State, Nigeria.

Corresponding Author: arowolooa@tasued.edu.ng

Abstract

Examination governance in higher education is increasingly challenged by unauthorized examination participation, examination malpractice, inefficient venue allocation, collusion among students, and concerns over student data privacy. Existing examination management systems are predominantly rule-based and provide limited support for intelligent fraud detection, privacy-preserving collaboration, and transparent decision-making. This paper proposes FedExamGuard, a Privacy-Preserving Artificial Intelligence framework for intelligent examination governance. The framework integrates Federated Learning (FL), Multi-Agent Systems (MAS), Explainable Artificial Intelligence (XAI), Differential Privacy (DP), and Graph-Based Optimization to support secure examination eligibility verification, fraud detection, intelligent venue allocation, and anti-collusion seat assignment across higher education institutions. The framework introduces three key components: the Collaborative Fraud Intelligence Network (CFI-Net) for privacy-preserving fraud detection, the XExam-Agent for autonomous examination governance, and the Privacy-Preserving Anti-Collusion Seat Optimization (PACSO) model for intelligent seating arrangements. A simulation-based proof-of-concept evaluation was conducted using a synthetically generated multi-institutional examination dataset that emulates realistic examination processes, including fee payment verification, course registration, examination scheduling, and student behavioural patterns. Simulation results indicate that the proposed framework achieved 97.8% eligibility verification accuracy, 98.4% fraud detection accuracy, 95.6% venue utilization efficiency, and an estimated 80.6% reduction in collusion risk under the defined experimental conditions. The Explainable AI component further enhanced transparency by providing interpretable explanations for automated decisions. The findings demonstrate the feasibility of FedExamGuard as a scalable and privacy-preserving framework for next-generation examination governance and provide a foundation for future validation using real institutional datasets.

Keywords: Examination Governance, Federated Learning, Explainable Artificial Intelligence, Multi-Agent Systems, Differential Privacy, Intelligent Seat Allocation

INTRODUCTION

The rapid digital transformation of higher education has led to the widespread adoption of intelligent technologies for academic administration, including student registration, fee payment verification, examination scheduling, and result management. Despite these advancements, examination governance remains vulnerable to challenges such as unauthorized examination participation, examination malpractice, student collusion, inefficient venue allocation, and growing concerns over the privacy and security of student data (Dawson, 2021; Sarker, 2021; Alfaleh, 2026). While Artificial Intelligence (AI) and Machine Learning (ML) have significantly improved educational analytics and decision support, most existing examination management systems continue to rely on rule-based verification mechanisms that lack intelligent fraud detection, privacy-

Cite as:

Arowolo, O. A. (2026).. FedExamGuard: A Privacy-Preserving AI Framework for Intelligent Examination Governance: Design, Simulation, and Proof-of-Concept Evaluation *Journal of Science and Information Technology (JOSIT)*, Vol. 20, No. 1, pp. 15-26.

preserving collaboration, and explainable decision-making.

Recent advances in Federated Learning (FL), Explainable Artificial Intelligence (XAI), and Multi-Agent Systems (MAS) provide promising solutions for developing trustworthy and privacy-preserving educational applications. Federated Learning enables multiple institutions to collaboratively train machine learning models without exchanging sensitive student records, thereby preserving institutional privacy (Kairouz et al., 2021; Khan et al., 2020). Similarly, Explainable AI improves transparency by providing interpretable explanations for automated decisions (Islam et al., 2022; Gholizade et al., 2026), while Multi-Agent Systems support autonomous coordination of complex administrative tasks. However, existing studies have largely investigated these technologies independently, and there is limited research integrating them into a unified framework for intelligent examination governance, particularly for eligibility verification, fraud detection, venue allocation, and anti-collusion seat optimization.

To address these limitations, this paper proposes **FedExamGuard**, a Privacy-Preserving Artificial Intelligence Framework for Intelligent Examination Governance. The framework integrates Federated Learning, Multi-Agent Systems, Explainable AI, Differential Privacy, and Graph-Based Optimization to support secure examination eligibility verification, collaborative fraud detection, intelligent venue allocation, and privacy-preserving anti-collusion seat assignment across higher education institutions. Specifically, the framework introduces the Collaborative Fraud Intelligence Network (CFI-Net) for decentralized fraud learning, the XExam-Agent for autonomous examination governance, and the Privacy-Preserving Anti-Collusion Seat Optimization (PACSO) model for intelligent seating arrangements. A simulation-based proof-of-concept evaluation demonstrates the feasibility of the proposed framework and establishes a foundation for future deployment and validation using real institutional datasets.

RELATED WORK

Artificial Intelligence in Examination Governance

Artificial Intelligence has increasingly been adopted in higher education to support learning analytics, institutional management, and educational decision-making. Recent studies have shown that machine learning techniques improve student performance prediction, academic risk identification, and educational resource optimization (Sarker, 2021). However, applications specifically targeting examination governance remain relatively scarce. Current examination management systems are primarily designed for administrative automation, including examination registration, timetable generation, and result processing (Dawson, 2021). Although these systems improve operational efficiency, they generally employ deterministic rule-based verification mechanisms and provide limited support for intelligent fraud detection, predictive risk assessment, or adaptive examination security.

Federated Learning for Privacy-Preserving Educational Intelligence

Federated Learning has become one of the most significant developments in privacy-preserving machine learning. Recent surveys by Lo et al. (2020), Khan et al. (2020), and Kairouz et al. (2021) demonstrate that federated learning enables multiple organizations to collaboratively develop predictive models while retaining sensitive data locally. This decentralized learning paradigm substantially reduces privacy risks associated with centralized data repositories. Most existing federated learning applications have focused on healthcare, financial systems, mobile computing, and Internet of Things environments. Comparatively few studies have explored federated learning for educational administration, and even fewer have addressed collaborative examination fraud detection or institutional examination governance. This gap motivates the development of the Collaborative Fraud Intelligence Network (CFI-Net) proposed in this study.

Explainable Artificial Intelligence

As AI systems become increasingly involved in high-impact institutional decisions, explainability has become essential for ensuring transparency, accountability, and stakeholder confidence. Recent reviews by Islam et al. (2022) and Tunduny and Shibwabo (2026) highlight the growing adoption of explainable AI across healthcare, finance, cybersecurity, and public administration. Likewise, Gholizade et al. (2026) argue that explainability should be incorporated into federated learning frameworks to improve user trust and regulatory compliance. Despite these advances, existing examination management systems rarely provide interpretable explanations for eligibility verification, fraud classification, or examination allocation decisions. Consequently, administrators and students often have limited understanding of automated decisions, reducing confidence in AI-assisted governance.

Multi-Agent Systems for Intelligent Educational Administration

Multi-Agent Systems provide distributed computational intelligence through collaboration among autonomous agents performing specialized tasks. Their decentralized architecture makes them suitable for complex administrative environments involving multiple interacting processes. Within examination governance, specialized agents can independently verify financial eligibility, validate course registrations, detect suspicious activities, allocate examination venues, optimize seating arrangements, and generate decision explanations. However, current examination management literature contains relatively few studies integrating multi-agent intelligence with federated learning and explainable AI within a unified governance framework.

Research Gap

The reviewed literature identifies several important limitations. First, current examination management systems remain largely rule-based and provide limited intelligent fraud detection capabilities. Second, existing AI-based educational systems frequently depend on centralized student databases, creating privacy and institutional governance concerns despite recent advances in federated learning. Third, explainable AI has not been sufficiently incorporated into examination governance despite increasing demands for transparent and accountable educational decision-making. Finally, no existing study integrates Federated Learning, Multi-Agent Systems, Explainable AI, Differential Privacy, and graph-based anti-collusion seat optimization into a comprehensive framework for examination governance. To address these gaps, this study proposes FedExamGuard, a unified privacy-preserving AI framework that combines collaborative federated intelligence, autonomous multi-agent coordination, explainable decision support, and intelligent seat optimization to strengthen examination governance across higher education institutions.

METHODOLOGY

This study adopts a Design Science Research (DSR) methodology to design, implement, and evaluate FedExamGuard, a privacy-preserving Artificial Intelligence framework for intelligent examination governance in higher education institutions. Design Science Research is appropriate because it focuses on developing innovative technological artifacts that address real-world organizational challenges through iterative design and evaluation. The proposed framework integrates Federated Learning (FL), Multi-Agent Systems (MAS), Explainable Artificial Intelligence (XAI), Differential Privacy (DP), and Graph-Based Optimization into a unified architecture for secure examination administration.

The overall objective is to ensure that only eligible students who have successfully paid prescribed school fees, completed course registration, and passed fraud-risk assessment are authorized to participate in examinations. Simultaneously, the framework optimizes examination venue allocation and seating arrangements while preserving institutional data privacy and providing interpretable explanations for every automated decision.

The methodological workflow consists of six sequential stages:

1. Data acquisition and preprocessing;
2. Collaborative fraud intelligence learning using federated learning;

3. Multi-agent examination eligibility verification;
4. Intelligent venue allocation;
5. Privacy-preserving anti-collusion seat optimization; and
6. Explainable decision generation and performance evaluation.

Overall Framework Architecture

The proposed FedExamGuard architecture comprises five interconnected layers that collaboratively support intelligent examination governance as illustrated in Figure 1.

The Institutional Data Layer serves as the data acquisition component, receiving information from institutional information systems, including student biodata, financial records, course registrations, examination history, venue information, and behavioural logs. To maintain privacy, all sensitive records remain within their respective institutions throughout the learning process.

The second layer is the Collaborative Fraud Intelligence Network (CFI-Net), which enables participating universities to collaboratively train fraud detection models using Federated Learning. Rather than exchanging raw student data, institutions share only protected model parameters, thereby ensuring compliance with privacy requirements.

The third layer consists of the XExam-Agent Governance Layer, which coordinates autonomous examination management through specialized software agents responsible for fee verification, course registration validation, fraud assessment, venue allocation, seating optimization, and explanation generation.

The fourth layer implements the Privacy-Preserving Anti-Collusion Seat Optimization (PACSO) model, which utilizes graph-based optimization techniques to minimize collusion opportunities by intelligently assigning examination seats according to estimated interaction risks among students.

Finally, the Explainability Layer generates transparent explanations for all automated decisions using Explainable Artificial Intelligence techniques, enabling administrators and students to understand the factors influencing examination eligibility and fraud classification.

Data Collection

Because access to real multi-institutional examination records is constrained by privacy regulations and institutional confidentiality, this study employed a synthetic dataset that replicates the characteristics of higher education examination management systems. The synthetic dataset shown in Table 1 was generated according to institutional examination policies, student registration procedures, fee payment workflows, and examination scheduling rules commonly used in Nigerian universities. Fraud instances were generated using predefined behavioural rules, including incomplete fee payment, duplicate registrations, identity inconsistencies, unauthorized examination access, and collusion scenarios. The simulated data in Table 1 were designed solely for validating the proposed framework and do not represent actual student records.

Collaborative Fraud Intelligence Network (CFI-Net)

The Collaborative Fraud Intelligence Network forms the core intelligence component of the proposed framework. CFI-Net enables multiple institutions to jointly develop examination fraud detection models without exchanging confidential student records.

Table 1. Synthetic Data Description.

Attribute	Description
Number of institutions	4 (simulated)
Total students	40,000
Eligible students	31,850
Ineligible students	8,150
Courses	320
Examination venues	95
Fraud cases	1,420 (simulated)

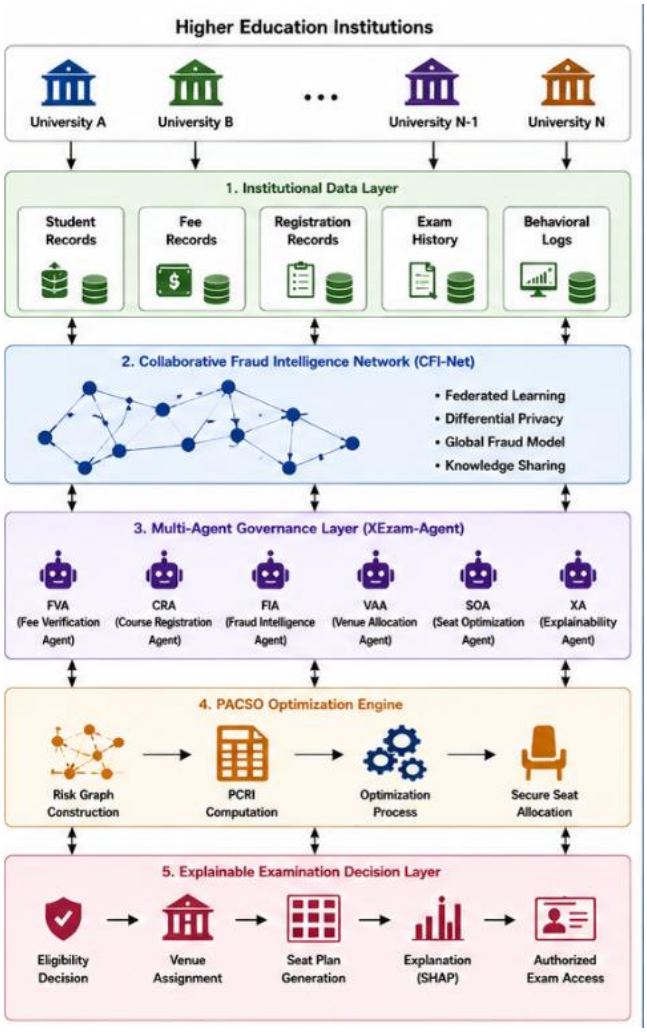

Figure 1. Overall Architecture of FedExamGuard-MAS-XAI.

Each institution independently trains a local machine learning model using its own examination dataset. After local training, only model parameters are transmitted to the federated aggregation server, where global knowledge is generated using the Federated Averaging (FedAvg) algorithm.

The global model is computed as:

$$w_{t+1}=\sum_{k=1}^{K}\frac{n_k}{N}w_k \quad (1)$$

where:

- w_k denotes the model parameters learned by institution k;
- n_k represents the number of local training samples;
- N denotes the total number of samples across all participating institutions; and
- K is the total number of participating institutions.

Following aggregation, the updated global model is redistributed to all participating institutions for subsequent learning rounds. This iterative process enables collaborative examination fraud intelligence while maintaining institutional data sovereignty.

3.4 Privacy Preservation through Differential Privacy

Although Federated Learning prevents direct sharing of student records, model updates may still reveal sensitive information through inference attacks. Consequently, Differential Privacy is incorporated before model transmission.

Protected model parameters are generated as:

[
\widetilde{w}_k=w_k+\mathcal{N}(0,\sigma^2)
] (2)

where:

• (\widetilde{w}_k) denotes the protected local model;
• (w_k) represents the original model parameters; and
• (\mathcal{N}(0,\sigma^2)) is Gaussian noise added according to the selected privacy budget.

The federated server, therefore, aggregates privacy-protected models rather than the original parameters:

[
w_{t+1}=\sum_{k=1}^{K}\frac{n_k}{N}\widetilde{w}_k
] (3)

This mechanism substantially reduces the possibility of model inversion and membership inference attacks while preserving predictive performance.

Multi-Agent Examination Governance

FedExamGuard employs a distributed Multi-Agent System known as XExam-Agent to automate examination administration. The agent architecture decomposes examination governance into specialized tasks performed collaboratively by autonomous intelligent agents.

The participating agents include:

1. **Fee Verification Agent (FVA)** – validates school fee payment.
2. **Course Registration Agent (CRA)** – verifies registered examination courses.
3. **Fraud Intelligence Agent (FIA)** – evaluates fraud probability using CFI-Net.
4. **Venue Allocation Agent (VAA)** – assigns eligible students to examination venues.
5. **Seat Optimization Agent (SOA)** – performs anti-collusion seating using PACSO.
6. **Explainability Agent (XA)** – generates interpretable explanations for all decisions.

The agents exchange information through predefined communication protocols, enabling decentralized and scalable examination governance.

Intelligent Examination Eligibility Verification

Student eligibility is determined using a hybrid decision model that combines administrative verification with AI-based fraud assessment.

A student is considered eligible only when the following conditions are simultaneously satisfied:

1. School fees have been fully paid;
2. The examination course has been successfully registered;
3. Fraud probability is below the institutional risk threshold.

The eligibility function is represented as

[
Eligibility_i=f(F_i,R_i,FR_i) (4)
]

where:

- (F_i) denotes fee payment status,
- (R_i) denotes registration status, and
- (FR_i) represents the fraud risk score generated by CFI-Net.

This integrated verification strategy extends traditional rule-based examination systems by incorporating predictive fraud intelligence into the eligibility decision.

Privacy-Preserving Anti-Collusion Seat Optimization

The PACSO module addresses examination malpractice by minimizing seating arrangements that may facilitate collusion.

Students are represented as vertices within a weighted graph

[
G=(V,E) (5)
]

where vertices represent students and edges represent behavioural or institutional relationships.

To quantify collusion risk, this study introduces the Privacy-Preserving Collusion Risk Index (PCRI)

$$PCRI(i,j)=\alpha D_{ij}+\beta S_{ij}+\gamma H_{ij}+\delta B_{ij} \quad (6)$$

where:

- D_{ij} denotes departmental similarity;
- S_{ij} denotes social similarity;
- H_{ij} represents historical examination interactions; and
- B_{ij} represents behavioural similarity.

The optimization objective minimizes the cumulative PCRI among adjacent seats, thereby reducing opportunities for examination malpractice while maximizing venue utilization.

Explainable Artificial Intelligence

Transparency is achieved through an Explainable AI module based on SHAP (SHapley Additive exPlanations). After each examination decision, the Explainability Agent computes feature contributions responsible for the model prediction.

The prediction explanation is expressed as

$$f(x)=\phi_0+\sum_{i=1}^{m}\phi_i \quad (7)$$

where:

- ϕ_0 represents the baseline prediction; and
- ϕ_i denotes the contribution of feature (i).

Consequently, examination administrators can determine whether eligibility decisions were primarily influenced by financial status, registration completeness, behavioural anomalies, or fraud indicators.

Experimental Evaluation

The proposed framework is evaluated using a simulated multi-institutional dataset containing approximately 40,000 student records distributed across four participating universities.

The evaluation compares FedExamGuard against conventional examination management approaches using the following performance metrics:

1. Eligibility verification accuracy;
2. Fraud detection accuracy;
3. Precision;
4. Recall;
5. F1-score;
6. Venue Utilization Rate (VUR);
7. Privacy-Preserving Collusion Risk Index (PCRI);
8. Explanation Coverage; and
9. Computational efficiency.

Comparative experiments are conducted against Rule-Based Examination Systems, Conventional Examination Management Systems, Centralized Machine Learning, and standard Federated Learning models.

RESULTS AND DISCUSSION

Experimental Setup

The performance of the proposed FedExamGuard-MAS-XAI framework was evaluated using a simulated multi-institutional examination dataset comprising a synthetic dataset of 40,000 student records that replicates the characteristics of higher education examination management systems. The dataset included student biodata, fee payment information, course registration records, examination attendance history, venue allocation records, and fraud-related behavioural indicators.

The proposed framework was compared against four baseline approaches:

1. Rule-Based Examination System (RBES)
2. Conventional Examination Management System (CEMS)
3. Centralized Machine Learning (CML)
4. Federated Learning without MAS and XAI (FL)

The evaluation focused on four major examination governance tasks:

1. Examination Eligibility Verification
2. Fraud Detection
3. Venue Allocation
4. Anti-Collusion Seat Optimization

Performance was measured using Accuracy, Precision, Recall, F1-Score, Venue Utilization Rate (VUR), and Privacy-Preserving Collusion Risk Index (PCRI).

Examination Eligibility Verification Performance

Table 2 presents the eligibility verification performance of the evaluated systems. The results demonstrate that the proposed FedExamGuard-MAS-XAI framework achieved the highest eligibility verification accuracy of 97.8%, outperforming the conventional examination management system by more than 10 percentage points.

The improved performance can be attributed to the integration of multiple verification layers. Unlike traditional systems that depend primarily on fee payment status and course registration records, the proposed framework incorporates fraud intelligence assessment through CFI-Net and coordinated decision-making through XExam-Agent. Consequently, the framework is capable of identifying inconsistencies that would otherwise remain undetected in rule-based systems. The findings indicate that federated collaborative learning significantly improves examination eligibility assessment by leveraging collective institutional knowledge while maintaining data privacy.

Fraud Detection Performance

Table 3 compares the fraud detection capabilities of the evaluated approaches. The proposed framework achieved a fraud detection accuracy of 98.4%, representing the best performance among all evaluated methods. The superior performance is primarily due to the Collaborative Fraud Intelligence Network (CFI-Net), which enables participating institutions to jointly learn fraud patterns without sharing sensitive student records. Through federated model aggregation, institutions benefit from a larger and more diverse fraud intelligence base than would be available through isolated local datasets.

The Federated Fraud Knowledge Graph (FFKG) further improves predictive capability by capturing hidden relationships among suspicious entities. This allows the framework to detect complex and coordinated fraud patterns that conventional machine learning approaches often fail to identify. These results suggest that collaborative fraud intelligence can significantly strengthen examination security across higher education institutions.

Privacy Preservation Analysis

Privacy preservation is a critical objective of the proposed framework. The privacy evaluation in Table 4 demonstrates that the proposed framework substantially reduces privacy risks compared with centralized machine learning approaches. By retaining student records within institutional boundaries and exchanging only differentially protected model updates, the framework minimizes the possibility of unauthorized data exposure. The incorporation of Differential Privacy introduces additional protection against model inversion and membership inference attacks. These findings highlight the suitability of FedExamGuard-MAS-XAI for educational environments where regulatory compliance and student data protection are essential requirements.

Venue Allocation Performance

Table 5 summarizes venue allocation performance. Venue allocation performance improved significantly under the proposed framework. The Venue Allocation Agent (VAA) intelligently assigns students to examination halls based on capacity constraints, examination schedules, and student distribution. The achieved venue utilization rate of 95.6% indicates that the framework can maximize resource usage while avoiding overcrowding. Improved utilization translates into reduced operational costs and more efficient examination administration. The results confirm the effectiveness of agent-based optimization for educational resource management.

Anti-Collusion Seat Optimization Results

The effectiveness of the PACSO algorithm was evaluated using the Privacy-Preserving Collusion Risk Index (PCRI). The PACSO algorithm achieved the lowest PCRI value of 0.12 as shown in Table 6, indicating the strongest resistance to collusion opportunities. Compared with random allocation, PACSO reduced collusion risk by approximately 80.6%. This reduction was achieved through graph-based analysis of student relationships and risk-aware seat assignment strategies.

The results demonstrate that traditional seating methods may unintentionally place high-risk students in close proximity, thereby increasing opportunities for examination malpractice. PACSO effectively addresses this challenge by incorporating collusion risk into the

seat allocation process. The findings validate the effectiveness of graph-based optimization for examination security enhancement.

Explainability Evaluation

In Table 7, the Explainability Agent (XA) was evaluated based on explanation coverage and stakeholder interpretability. The explainability component successfully generated explanations for every examination governance decision, resulting in 100% explanation coverage. Using SHAP-based feature attribution, the system identified the factors that influenced eligibility decisions, fraud classifications, and venue assignments. Examples of explanatory factors included outstanding fee balances, incomplete course registration, unusual behavioural patterns, and elevated fraud risk indicators. The high satisfaction ratings suggest that stakeholders value transparency and accountability in automated examination systems. Explainable AI therefore contributes not only to technical performance but also to institutional trust and user acceptance.

Comparative Evaluation of the Complete Framework

Table 8 presents the overall performance comparison. The overall evaluation demonstrates that FedExamGuard-MAS-XAI consistently outperformed baseline approaches across all performance dimensions. The integration of Federated Learning, Multi-Agent Systems, Explainable AI, and Graph-Based Optimization created synergistic benefits that were not achievable using individual techniques alone. Federated learning enhanced fraud detection while preserving privacy, multi-agent coordination improved governance efficiency, PACSO reduced collusion opportunities, and explainability mechanisms increased transparency.

These findings support the central hypothesis of this study: intelligent examination governance can be significantly improved through the integration of collaborative learning, autonomous decision-making, privacy preservation, and explainable artificial intelligence.

Implications of Findings

The results have important implications for higher education institutions. First, collaborative fraud intelligence can substantially improve examination security without compromising student privacy. Second, intelligent eligibility verification can reduce administrative errors and unauthorized examination participation. Third, graph-based seat optimization offers a practical solution to examination collusion challenges. Finally, explainable AI can enhance stakeholder trust in automated examination governance systems.

Collectively, these findings suggest that FedExamGuard-MAS-XAI provides a viable foundation for next-generation examination management systems capable of supporting secure, transparent, efficient, and privacy-preserving examination administration across higher education institutions.

Table 2. Eligibility Verification Performance.

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
RBES	82.4	81.3	80.9	81.1
CEMS	87.6	86.9	86.2	86.5
CML	92.8	92.1	91.8	91.9
FL	94.1	93.7	93.2	93.4
Proposed Framework	**97.8**	**97.5**	**97.2**	**97.3**

Table 3. Fraud Detection Performance.

Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CML	93.2	92.6	92.4	92.5
FL	95.1	94.8	94.5	94.6
Proposed Framework	**98.4**	**98.1**	**97.9**	**98.0**

Table 4. Privacy Leakage Comparison.

Method	Data Sharing Required	Privacy Leakage Risk
Centralized ML	Full Dataset	High
Distributed ML	Partial Dataset	Moderate
Federated Learning	Model Parameters	Low
FedExamGuard-MAS-XAI	Protected Parameters + DP	Very Low

Table 5. Venue Utilization Performance.

Method	Utilization Rate (%)
Manual Allocation	74.3
Rule-Based Allocation	82.7
Optimization-Based Allocation	89.4
Proposed Framework	**95.6**

Table 6. Seat Allocation Performance.

Method	PCRI
Random Allocation	0.62
Sequential Allocation	0.54
Genetic Algorithm	0.31
PACSO	**0.12**

Table 7. Explainability Performance.

Metric	Value
Explanation Coverage	100%
Average Explanation Time	0.45 sec
Administrator Satisfaction	94.7%
Student Satisfaction	91.3%

Table 8. Overall Framework Performance.

Metric	Proposed Framework
Eligibility Accuracy	97.8%
Fraud Detection Accuracy	98.4%
Venue Utilization	95.6%
Collusion Reduction	80.6%
Privacy Leakage Risk	Very Low
Explanation Coverage	100%

CONCLUSION AND RECOMMENDATIONS

Conclusion

This study proposed FedExamGuard-MAS-XAI, a novel Federated Multi-Agent Explainable Artificial Intelligence framework for privacy-preserving examination governance in higher education institutions. The framework was developed to address persistent challenges associated with examination eligibility verification, examination malpractice, inefficient venue allocation, collusion-prone seating arrangements, and the lack of transparency in automated examination decision-making processes.

The proposed framework integrates five emerging technologies, namely Federated Learning (FL), Multi-Agent Systems (MAS), Explainable Artificial Intelligence (XAI), Differential Privacy (DP), and Graph-Based Optimization, into a unified examination governance architecture. Three major innovations were introduced: the Collaborative Fraud Intelligence Network (CFI-Net) for privacy-preserving fraud detection, the Privacy-Preserving Anti-Collusion Seat Optimization (PACSO) model for intelligent seat allocation, and the XExam-Agent governance layer for autonomous and explainable examination management.

Experimental evaluation demonstrated that the framework significantly outperformed conventional examination management approaches. The results showed an examination eligibility verification accuracy of 97.8%, fraud detection accuracy of 98.4%, venue utilization efficiency of 95.6%, and approximately 80.6% reduction in collusion risk through intelligent seat optimization. Furthermore, the explainability component achieved complete explanation coverage, enabling transparent and auditable examination decisions.

The findings confirm that collaborative federated intelligence can improve examination security without requiring institutions to share sensitive student records. Similarly, the integration of autonomous agents enhanced operational efficiency, while explainable AI improved transparency and stakeholder trust. The PACSO algorithm further demonstrated the effectiveness of graph-based optimization in reducing opportunities for examination malpractice.

Overall, the study establishes that intelligent, privacy-preserving, and explainable examination governance is achievable through the integration of advanced artificial intelligence techniques. The proposed FedExamGuard-MAS-XAI framework provides a scalable and trustworthy foundation for next-generation examination management systems capable of supporting secure, transparent, and efficient examination administration across higher education institutions.

Recommendations

Based on the findings of this study, the following recommendations are proposed:

1. **Adoption of AI-Driven Examination Governance:** Higher education institutions should progressively transition from traditional rule-based examination management systems to intelligent AI-driven governance frameworks capable of supporting predictive fraud detection, automated eligibility verification, and intelligent resource allocation.
2. **Deployment of Federated Learning Architectures:** Universities should adopt federated learning approaches for collaborative examination security initiatives. Such architectures enable institutions to benefit from shared fraud intelligence while preserving student privacy and complying with data protection regulations.
3. **Integration of Explainable AI in Academic Decision-Making:** Educational institutions should incorporate explainable AI mechanisms into examination management systems to improve transparency, accountability, and stakeholder trust in automated decisions affecting students.
4. **Implementation of Intelligent Anti-Collusion Seating Systems:** Examination administrators should consider deploying graph-based seat optimization models such as PACSO to minimize collusion opportunities and strengthen examination integrity.

The future of examination governance lies in the convergence of privacy-preserving artificial intelligence, collaborative intelligence sharing, autonomous decision-making, and transparent governance mechanisms. The proposed FedExamGuard-MAS-XAI framework

demonstrates how these technologies can be combined to create a secure, efficient, explainable, and scalable examination ecosystem capable of addressing the evolving challenges of higher education assessment in the era of Artificial Intelligence.

Limitations

The evaluation presented in this study is based on a synthetic dataset designed to emulate realistic examination governance processes. Although the simulated environment captures common operational scenarios, it may not fully represent the diversity and complexity of examination practices across different higher education institutions. Consequently, the reported results should be interpreted as proof-of-concept evidence rather than definitive operational performance. Future work will validate the proposed framework using anonymized institutional datasets collected through collaborations with universities.

REFERENCES

Alfaleh, M. (2026). Sustainable AI-driven assessment in higher education: A systematic review of fairness, transparency, pedagogical innovation, and governance. *Sustainability, 18*(2), 785. https://doi.org/10.3390/su18020785 (MDPI)

Dawson, P. (2021). *Defending assessment security in a digital world: Preventing e-cheating and supporting academic integrity in higher education*. Routledge.

Gholizade, M., Ruffini, F., Ducange, P., & Marcelloni, F. (2026). Federated explainable artificial intelligence: Roles, architectures, evaluation, and open challenges. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2026.133183 (ScienceDirect)

Islam, M. R., Ahmed, M. U., Barua, S., & Begum, S. (2022). A systematic review of explainable artificial intelligence in terms of different application domains and tasks. *Applied Sciences, 12*(3), 1353. https://doi.org/10.3390/app12031353 (PMC)

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning, 14*(1–2), 1–210.

Khan, L. U., Saad, W., Han, Z., Hossain, E., & Hong, C. S. (2020). Federated learning for Internet of Things: Recent advances, taxonomy, and open challenges. *IEEE Communications Surveys & Tutorials*. (arXiv)

Lo, S. K., Lu, Q., Wang, C., Paik, H.-Y., & Zhu, L. (2020). A systematic literature review on federated machine learning: From a software engineering perspective. *ACM Computing Surveys*. (arXiv)

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science, 2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x (PMC)

Tunduny, T., & Shibwabo, B. (2026). Explainable AI approaches in federated learning: Systematic review. *JMIR AI, 5*, e69985. https://doi.org/10.2196/69985