# PROCTORING AT A NEW STAGE: NEXT-GENERATION AI SYSTEMS FOR ENSURING TRUST IN DIGITAL EXAMINATIONS

Mo'minov Haydar Muhiddinovich
*CEO, Univora IT Group LLC | Proctoring.uz*
University of Tashkent for Applied Sciences (UTAS), Tashkent, Uzbekistan
E-mail: univorarsa@gmail.com / haydarmuminov@gmail.com / | proctoring.uz

Abstract. The rapid proliferation of online and hybrid learning environments, accelerated by institutional digitalisation programmes across Central Asia and beyond, has elevated academic integrity to a critical challenge worldwide. Legacy proctoring solutions — whether in-person supervision or first-generation remote monitoring software — prove systematically inadequate against sophisticated cheating strategies enabled by generative artificial intelligence (AI) tools. This paper presents a rigorous analysis of a next-generation AI-powered proctoring architecture deployed at scale, combining four integrated layers: (1) multi-modal behavioural analysis (facial landmark tracking, gaze estimation, head pose, keyboard dynamics); (2) large language model (LLM) semantic anomaly detection and stylometric authorship verification; (3) federated, differential-privacy-preserving learning ($\varepsilon = 0.5$) that ensures no raw biometric data leaves institutional premises; and (4) explainable AI (XAI) decision support via SHAP values, producing human-readable integrity reports with full audit trails. Drawing on real-world deployment data from Proctoring.uz — the first locally developed AI proctoring platform in Central Asia, covering 129,463 examination sessions across three government ministries and multiple higher education institutions in Uzbekistan — we demonstrate that the proposed hybrid architecture achieves a detection accuracy of 99.8%, an F1-score of 0.998, and a false positive rate of 0.2%, while reducing average review time per flagged session from 42 minutes to 11 minutes relative to legacy systems. The platform is formally certified by the Uzbek Ministry of Justice (Certificate No. DGU 37103), recognised as a top-10 winner among nearly 300 startups at the U-Solve Incubation Programme (supported by UK International Development / VirtualAccelerate), and selected for presentation at GITEX AI Central Asia 2026 (Almaty, Kazakhstan). A structured competitive evaluation against ProctorU, Honorlock, and Proctorio, based on publicly documented capabilities, demonstrates 100% coverage of 18 monitoring parameters with a price point estimated at 10–15 times lower than international alternatives. Ethical dimensions — algorithmic bias, surveillance governance, and explainability limits — are examined alongside technical results.

Keywords: *AI proctoring; digital examinations; academic integrity; federated learning; explainable AI (XAI); behavioural biometrics; generative AI threats; deepfake detection; SHAP; trusted assessment; Central Asia EdTech; data sovereignty.*

## 1. Introduction

The global expansion of remote and hybrid learning — accelerated by the COVID-19 pandemic and sustained by national digitalisation strategies including Uzbekistan's "Digital Uzbekistan 2030" programme — has fundamentally restructured academic assessment. According to UNESCO (2023), more than half of higher education institutions worldwide had adopted some form of remote examination infrastructure by 2023, a proportion that continues to grow. This

structural shift has introduced an asymmetric security problem: examination delivery has become digital, yet the integrity mechanisms protecting that delivery have not evolved at the same pace.

The public availability of large language models (LLMs) with extended reasoning, multi-modal understanding, and near-real-time response latency has rendered many traditional integrity guarantees obsolete. Legacy proctoring systems — relying on webcam streams, screen recording, and coarse keystroke logging — are now routinely circumvented through secondary devices, AI-generated answer delivery, adversarial prompt injection against AI-based grading pipelines, and deepfake identity substitution. The release of multi-modal frontier models since mid-2023 has widened the detection gap further and at accelerating pace.

Proctoring.uz was designed to close this gap. As the first locally developed AI proctoring platform in Central Asia, it has to date processed 129,463 real-world examination sessions for three government ministries and multiple universities in Uzbekistan, achieving 99.8% detection accuracy across 18 monitoring parameters — parameters that global competitors such as ProctorU, Honorlock, and Proctorio do not all offer. The platform is formally certified by the Uzbek Ministry of Justice (Certificate No. DGU 37103), validated as a top-10 winner at the internationally judged U-Solve Incubation Programme, and selected for presentation at GITEX AI Central Asia 2026 in Almaty, Kazakhstan.

This paper makes four contributions:

- a taxonomy of AI-assisted cheating strategies and their detectability by legacy versus next-generation systems (Section 2);
- a formally specified four-layer proctoring architecture integrating multi-modal behavioural analysis, semantic anomaly detection, federated learning, and explainable decision support (Section 3);
- empirical performance evaluation across 129,463 sessions, including detection accuracy, false positive rates, F1-score, review efficiency, and student acceptance (Section 5); and
- a structured ethical analysis of bias, surveillance governance, and algorithmic accountability (Section 6).

## 2. Threat Landscape: Generative AI and Academic Dishonesty

### 2.1 Taxonomy of AI-Enabled Academic Dishonesty

The introduction of frontier LLMs has fundamentally altered the risk profile of digital examinations. We identify six primary categories of AI-assisted academic dishonesty in the 2024–2026 period:

- Direct answer generation. A student submits examination questions to an external LLM service via a secondary device, browser tab, or wearable, and presents the response as original work.
- Paraphrase laundering. AI-generated content is processed through paraphrasing tools to reduce stylometric similarity to known LLM outputs, bypassing textual similarity detectors while preserving semantic content.
- Adversarial prompt injection. Malicious instructions embedded within submitted responses or file uploads aim to manipulate AI-assisted grading systems — a systemic risk as automated marking becomes more prevalent.
- Multi-modal wearable assistance. Vision-language models operating outside the proctoring environment — via smart glasses or hidden earpieces — provide real-time guidance on hardware not monitored by proctoring software.

- Deepfake identity substitution. A student is replaced by a confederate via a real-time face-replacement deepfake injected into the webcam feed. Advances in generative video since 2023 have made high-quality live deepfake injection accessible to non-expert actors.
- Collaborative answer networks. Structured peer groups distribute examination questions via encrypted messaging platforms immediately after session commencement, with AI-assisted answers shared back in near-real-time.

### 2.2 Empirical Evidence of AI-Assisted Cheating

Multiple independent surveys confirm the growing prevalence of AI-assisted academic dishonesty. The Wiley Academic Integrity Survey (2024), conducted with 850 instructors and 2,067 students across the United States and Canada, found that 96% of instructors believed at least some of their students had cheated in the past year — a substantial increase from 72% in 2021 [2]. A 2024 survey by Inside Higher Ed found that 45% of students used AI in their classes during the previous academic year [3]. Research using indirect list-experiment methods, which reduce social desirability bias, estimated that approximately 22% of UK university students admitted to cheating using generative AI during the 2023–2024 academic year [4].

Real-world deployment data from Proctoring.uz corroborates and extends these survey findings. Across 129,463 monitored examination sessions, the platform's detection layers flagged the incident distribution presented in Table 2 (Section 5.1), providing one of the first large-scale empirical datasets of AI-assisted cheating incidence in a Central Asian deployment context.

### 2.3 Detection Effectiveness: Legacy Versus Next-Generation Systems

Table 1 presents a structured comparison of detection effectiveness across cheating strategies for legacy single-signal systems and the proposed next-generation architecture, based on evaluation against the 1,240-session ground-truth dataset described in Section 5. Detection rates for the proposed system reflect internal evaluation results. Legacy detection ranges are indicative estimates derived from published literature and vendor-reported benchmarks [7].

*Table 1. Cheating strategy taxonomy and comparative detection effectiveness: legacy systems versus proposed next-generation architecture*

| Cheating Strategy | Legacy Detection Rate | Next-Gen AI Detection Rate | Primary Mitigation Mechanism |
|---|---|---|---|
| Direct LLM answer generation | 15–30% | 88–99.8% | Semantic anomaly detection; behavioural baseline deviation scoring |
| Paraphrase laundering | < 10% | 65–78% | Authorship verification; keyboard dynamics analysis |
| Secondary device / external screen | 55–70% | 90–96% | Real-time gaze tracking; head pose estimation via PnP-RANSAC |
| Wearable AI assistance | < 5% | 60–72% | Biometric anomaly profiling; keyboard dynamics fingerprinting |
| Adversarial prompt | None | 85–91% | Grading sandbox isolation; input |

| | | | |
|---|---|---|---|
| injection | (0%) | | sanitisation pipeline |
| Deepfake identity substitution | None (0%) | 99.8% | 3D liveness detection; frame-level 68-point facial model analysis |

Note. *Detection rates for the Proctoring.uz system are based on internal evaluation against a manually verified ground-truth dataset (n = 1,240 sessions; Cohen's κ = 0.91). Legacy detection ranges are indicative estimates drawn from published academic literature and publicly available vendor documentation [7]. These estimates are not directly comparable across platforms, as evaluation methodologies differ.*

The data in Table 1 reveal a systematic performance gap, particularly for modern high-impact attack vectors. Legacy systems exhibit near-zero detection capability against deepfake identity substitution and adversarial prompt injection — precisely the attack vectors most likely to be exploited by technically sophisticated adversaries. The proposed architecture addresses these gaps through multi-layer design rather than incremental tuning of existing approaches.

## 3. Architectural Framework for Next-Generation AI Proctoring

The proposed architecture consists of four interdependent layers, each addressing a distinct threat category. The layers are designed to be privacy-preserving by construction, explainable by default, and scalable across contexts ranging from single classrooms to national examination systems.

### 3.1 Layer 1: Multi-Modal Behavioural Analysis

The foundation is a continuous monitoring pipeline fusing five data streams:

- Facial landmark tracking (68-point model, 30 fps): Continuous detection of facial keypoints enabling precise head position characterisation, micro-expression sequencing, and three-dimensional face geometry required for liveness verification and deepfake detection.
- Gaze estimation via appearance-based CNN: Estimates the student's point of visual attention at sub-second resolution. Sustained off-screen gaze is flagged as a gaze-deviation violation.
- Head pose estimation via PnP-RANSAC: Derives three-dimensional orientation (yaw, pitch, roll) from 2D landmark projections. Head rotations inconsistent with forward-facing engagement, correlated with gaze deviations, constitute strong anomaly signals.
- Keyboard dynamics (dwell time, flight time, digraph latency): Characterises the temporal micro-structure of typing. Anomalous bursts consistent with copy-paste from an LLM output are detectable through deviation from the baseline.
- Mouse movement entropy: Scripted or tool-assisted cursor movement produces entropy signatures distinct from natural human navigation.

Each stream is modelled independently during a 5–10 minute baseline calibration period prior to examination commencement. Deviations are scored using an isolation forest anomaly detector trained on a federated dataset of over 180,000 examination sessions. The system computes a real-time integrity score $I(t) \in [0, 1]$ updated every second, where higher values indicate greater deviation from the student-specific behavioural baseline.

### 3.2 Layer 2: LLM-Based Semantic Anomaly Detection

The second layer addresses the semantic integrity of submitted written content — a dimension that purely behavioural systems cannot evaluate. The system applies a fine-tuned classification model trained to identify three statistical signatures characteristic of LLM-generated text:

- Perplexity uniformity: LLM outputs exhibit unusually low variance in per-token perplexity. Human-authored text shows characteristically higher perplexity variation, particularly under examination time pressure.
- Sentence-length burstiness: Human writing exhibits power-law distributed sentence length variation. LLM outputs show statistically more uniform length distributions that are detectable even after surface-level paraphrasing.
- Semantic coherence pattern analysis: LLMs produce globally coherent responses uncharacteristic of improvised examination writing. The system identifies coherence patterns inconsistent with the student's established writing profile.

Authorship verification cross-references the stylometric fingerprint of examination responses against a corpus of verified writing samples from across the academic term. A Bayesian probability model aggregates the AI generation likelihood with the authorship distance metric to produce a composite content integrity score $C \in [0, 1]$ per submitted response.

### 3.3 Layer 3: Federated and Privacy-Preserving Learning

A fundamental limitation of centralised proctoring architectures is the transmission of raw video, audio, and biometric data to third-party servers, creating substantive privacy risks and compliance challenges under Uzbekistan's Law on Personal Data (2019), GDPR, and the OECD AI Principles [8, 9]. Proctoring.uz operates on federated learning principles [5]: institutional nodes train local anomaly detection models on their own examination data; only gradient updates — not raw data — are transmitted to a central aggregation server.

Three formal privacy mechanisms are applied in combination:

- Differential privacy ($\varepsilon = 0.5$): Calibrated Gaussian noise applied to all gradient updates before transmission, providing a mathematically guaranteed upper bound on the information that can be inferred about any individual record from the shared model parameters.
- Secure multi-party computation (SMC): Gradient aggregation via SMC protocols ensures that no single party — including the aggregation server — can observe the unencrypted gradient contributions of any individual institution.
- On-premise data residency: Video streams, audio data, and biometric measurements never leave institutional premises. This architecture satisfies Uzbekistan's data localisation requirements applicable to government examination systems.

### 3.4 Layer 4: Explainable and Auditable Decision Support

SHAP (SHapley Additive exPlanations) values [6] are computed across the joint feature space of behavioural and semantic inputs for every flagged session, identifying the specific signals — and their relative contributions — that drove an elevated integrity score. These values are rendered as a human-readable integrity report accessible to designated human reviewers, faculty, and — upon request — the examined student.

All analytical outputs are logged in an immutable, time-stamped audit trail. Students retain an unconditional right to request a full explanation for any flag. All final adjudication decisions — including any disciplinary consequences — are made exclusively by designated human reviewers. The AI architecture operates as a decision-support tool: it identifies, quantifies, and explains anomalies; it does not determine outcomes. This distinction is architecturally enforced through the system design.

## 4. Competitive Analysis and Market Positioning

Table 3 presents a structured evaluation of Proctoring.uz against ProctorU, Honorlock, and Proctorio — three platforms with the largest international market presence. Feature coverage for competitor platforms is assessed by the author against publicly available product documentation (2024–2025), as none of these platforms publish formal feature coverage benchmarks against a common standard. Price estimates for competitor platforms reflect indicative third-party analyst ranges; all three platforms operate on bespoke institutional pricing and do not publish list prices.

*Table 3. Competitive evaluation: Proctoring.uz versus international proctoring platforms (based on publicly documented capabilities)*

| Feature / Criterion | Proctoring.uz | ProctorU | Honorlock | Proctorio |
|---|---|---|---|---|
| Monitoring parameters covered (of 18, author evaluation) | 18/18 (100%) | ~10/18 | ~12/18 | ~9/18 |
| Estimated price per session (indicative) | $0.50–1.50 | Not published | Not published | Not published |
| Deepfake identity detection | ✓ | ✗ | ✗ | ✗ |
| Mobile device blocking | ✓ | ✗ | ✗ | ✗ |
| Remote Desktop / VM protection | ✓ | ✗ | ✗ | ✗ |
| Local servers (data sovereignty) | ✓ | ✗ | ✗ | ✗ |
| Local language interface (Uzbek/Russian) | ✓ | ✗ | ✗ | ✗ |
| Government regulatory registration | ✓ DGU 37103 | ✗ | ✗ | ✗ |

Note. ✓ = *feature present and documented;* ✗ = *feature absent from publicly available documentation. Coverage fractions (~10/18, ~12/18, ~9/18) represent the author's evaluation of competitor capabilities against Proctoring.uz's 18-parameter framework and should be understood as indicative rather than independently verified. Competitor pricing is not publicly disclosed; the price range for Proctoring.uz ($0.50–1.50 per session) is the platform's own institutional pricing. The "10–15×" price advantage statement is based on third-party market reports comparing similar per-session proctoring costs in the region.*

The evaluation reveals a distinctive differentiation profile. Proctoring.uz is the only evaluated platform with documented deepfake identity detection, mobile device blocking, remote desktop protection, and full local data sovereignty — capabilities that address compliance requirements unmet by international alternatives, particularly for government-sector deployments subject to national data localisation frameworks.

## 5. Experimental Evaluation

### 5.1 Dataset and Evaluation Protocol

The proposed architecture was evaluated on the production deployment dataset of 129,463 real-world examination sessions conducted between 2024 and 2025 across three government ministries and multiple higher education institutions in Uzbekistan. Ground-truth labels were established by three independent expert reviewers who manually verified a stratified random sample of 1,240 sessions (0.96% of the total dataset) for the presence or absence of academic dishonesty. Inter-rater agreement (Cohen's κ) across reviewers was 0.91, indicating near-perfect reliability. All performance metrics reported in Section 5.2 refer to this ground-truth evaluation dataset unless otherwise stated.

Incident distribution across the full 129,463-session production dataset is summarised in Table 2. These figures reflect flagged events prior to human review; not all flagged events resulted in confirmed integrity findings.

*Table 2. Incident distribution across 129,463 production examination sessions (flagged events, prior to human adjudication)*

| Incident Category | Events Detected | % of Total Sessions | Architectural Layer Responsible |
|---|---|---|---|
| Voice / conversation incidents | 278,080 | 214.8%* | Layer 1 — Audio stream anomaly detection |
| Face-mismatch events | 59,610 | 46.1% | Layer 1 — 3D liveness + deepfake module |
| Gaze-deviation violations | 45,310 | 35.0% | Layer 1 — Appearance-based CNN gaze estimator |
| Full-screen exit events | 42,970 | 33.2% | Layer 1 — Screen integrity monitor |
| Focus-change anomalies | 26,110 | 20.2% | Layer 1 — Keyboard / mouse entropy analyser |

Note. **Voice / conversation incidents per session can exceed 100% because multiple events may be flagged within a single session. All flagged events are subject to human reviewer assessment before any determination of academic dishonesty is made.*

## 5.2 Detection Performance Metrics

Performance metrics computed against the 1,240-session ground-truth evaluation dataset are presented in Table 4. Figures for legacy single-signal systems reflect ranges reported in published academic literature and vendor documentation; methodologies differ across sources and the comparison should be understood as indicative.

*Table 4. Detection performance: Proctoring.uz next-generation system versus reported legacy system baseline*

| Performance Metric | Proctoring.uz Next-Generation System | Legacy Single-Signal Systems (Reported Baseline) |
|---|---|---|
| Detection accuracy (Recall) | 99.8% | 61–72% (published range) |

| | | |
|---|---|---|
| False positive rate | 0.2% | 8.7% (published range) |
| F1-score | 0.998 | 0.68–0.74 (published range) |
| Average review time per flagged session | 11 min | 42 min |
| Student complaint rate | 0.2% | 14.3% |
| Privacy-acceptable rating (student survey) | 78% (internal survey) | Not reported |

Note. *Recall, false positive rate, and F1-score computed against 1,240 manually verified ground-truth sessions (Cohen's κ = 0.91). Review time and complaint rate computed across the full 129,463-session production dataset. Legacy system figures are indicative ranges from published academic literature and vendor-reported benchmarks; they do not represent a controlled head-to-head evaluation on the same dataset.*

The proposed architecture achieves a detection accuracy (Recall) of 99.8% with a false positive rate of 0.2% — a 97.7% relative reduction compared to the legacy baseline range. Average flagged-session review time decreases from 42 minutes to 11 minutes, representing a 73.8% efficiency improvement. Student complaint rates fall from 14.3% to 0.2%. These metrics are derived from the platform's internal evaluation and have not been independently audited by a third-party laboratory, a limitation to be addressed in future work.

### 5.3 Privacy and Compliance Evaluation

Independent legal review confirmed that the federated architecture satisfies Uzbekistan's Law on Personal Data (2019) and OECD AI Principles (2024) [8, 9]. Formal privacy auditing verified the differential privacy guarantee at ε = 0.5. No raw biometric data left institutional premises during the evaluation period. Student post-examination surveys (conducted internally across a subset of examination sessions) indicated that 78% of respondents rated the federated approach as significantly more acceptable from a privacy standpoint compared to conventional cloud-based proctoring.

### 5.4 Market Validation and International Recognition

Proctoring.uz was selected as one of the top 10 projects from 20 competing startups at the U-Solve Incubation Programme, a competitive accelerator supported by UK International Development, VirtualAccelerate, Yoshlar Ventures, Intro Group, and U-Enter Innovation Centre, assessed by a jury including IT Park Ventures, SQB Ventures, United Ventures, and UzVC. This earned an invitation to present at GITEX AI Central Asia 2026 (Almaty, Kazakhstan, 4–5 May 2026). The platform holds formal Ministry of Justice certification (DGU 37103) and is actively deployed across three government ministries and multiple universities in Uzbekistan.

## 6. Ethical Considerations and Limitations

### 6.1 Algorithmic Bias and Fairness

Behavioural models trained in specific institutional or cultural contexts may exhibit performance disparities across demographic groups. Research in the facial recognition literature has documented systematic accuracy differentials, and analogous effects are plausible in gaze and keyboard-dynamics models across Uzbekistan's multi-ethnic population. The federated architecture provides some structural protection against distributional mismatch, but does not

eliminate it. Continuous bias auditing with disaggregated demographic reporting is maintained across all deployment sites and is a non-negotiable operational requirement.

### 6.2 Surveillance Governance and Mission Creep

The availability of high-quality behavioural monitoring infrastructure creates incentives to extend its use beyond examination contexts. Technical controls — including cryptographically enforced time-bounded activation and institutionally governed data retention policies — enforce strict use-case limitation. These controls are implemented in the platform and are subject to governance audit.

### 6.3 Explainability Limits and Algorithmic Accountability

Despite the SHAP-based XAI layer, deep neural network representations cannot be fully translated into non-expert terms in all cases. The architectural principle that AI provides decision support while humans make adjudication decisions provides a procedural safeguard that remains valid in the presence of partial explainability limitations. Continued investment in interpretability research is required.

### 6.4 Limitations of the Present Study

The performance metrics reported in Section 5 are based on the platform's internal evaluation; no independent third-party laboratory validation has yet been conducted. The competitive feature comparison in Section 4 is based on publicly available competitor documentation and the author's assessment — it has not been verified by the competitor platforms. The student privacy survey was conducted internally, not by an independent research team. These limitations constrain the generalisability of the findings and will be addressed in future peer-reviewed evaluation.

## 7. Conclusions

This paper has presented a comprehensive next-generation AI proctoring architecture and evaluated it through real-world deployment data from Proctoring.uz — the first locally developed AI proctoring platform in Central Asia.

The four-layer system combining multi-modal behavioural analysis, LLM-based semantic anomaly detection, federated learning ($\varepsilon = 0.5$), and SHAP-based explainable decision support achieves, according to internal evaluation, a detection accuracy of 99.8%, an F1-score of 0.998, and a false positive rate of 0.2% across 129,463 real-world examination sessions. Four conclusions follow:

- A paradigm shift is required. Effective response to generative AI threats requires a transition from reactive surveillance to proactive, probabilistic, multi-modal integrity assurance.
- Privacy and performance are complementary. Federated learning with differential privacy ($\varepsilon = 0.5$) enables high-accuracy anomaly detection without transmitting raw biometric data, demonstrating that data sovereignty and detection performance can be jointly optimised.
- Explainability and human oversight are non-negotiable. Architecturally enforced human adjudication — supported by SHAP-based explanations and immutable audit trails — is both ethically necessary and practically feasible at scale.
- Locally developed platforms can achieve global competitiveness. The Proctoring.uz deployment demonstrates that an emerging-market EdTech platform can achieve broad functional coverage, international recognition, and cost advantages over established global competitors through architecture-first design tailored to local regulatory requirements.

Future work will address: (i) independent third-party laboratory validation of performance metrics; (ii) adversarial robustness testing against adaptive cheating strategies; (iii) longitudinal bias monitoring across demographic groups; and (iv) development of open-source reference implementations to support under-resourced institutions.

*As generative AI continues to evolve, the systems designed to ensure that educational credentials reflect genuine competence must evolve in parallel — treating students not as suspects to be surveilled, but as participants in a shared commitment to the integrity of learning.*

Acknowledgements

The author gratefully acknowledges the institutional partnership between the University of Tashkent for Applied Sciences (UTAS) and Univora IT Group LLC / RealSoft Academy MChJ, formalised through a Memorandum of Understanding. The 5+1 Internship Programme — under which UTAS students undertake one day per week of supervised practice at Univora IT Group / RealSoft Academy under industry mentorship — has contributed to the platform's real-world evaluation through joint examination session monitoring. The author also thanks the U-Solve Incubation Programme (UK International Development / VirtualAccelerate), the evaluating jury, and the organising committee of GITEX AI Central Asia 2026 for independent validation of the platform's merits.

References

[1] UNESCO. (2023). Digital learning and education: Trends, opportunities and challenges. UNESCO Publishing. https://www.unesco.org/en/digital-education

[2] Wiley. (2024). The Latest Insights into Academic Integrity: Instructor & Student Experiences, Attitudes, and the Impact of AI 2024 Update. John Wiley & Sons. https://newsroom.wiley.com/press-releases/press-release-details/2024/AI-Has-Hurt-Academic-Integrity-in-College-Courses/

[3] Inside Higher Ed. (2024, July 29). Students and professors expect more cheating thanks to AI. Inside Higher Ed. https://www.insidehighered.com/news/tech-innovation/artificial-intelligence/2024/07/29/students-and-professors-expect-more

[4] Lancaster, T., & Bishop, A. (2025). How vulnerable are UK universities to cheating with new GenAI tools? A pragmatic risk assessment. Assessment & Evaluation in Higher Education. https://doi.org/10.1080/02602938.2025.2511794

[5] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS 2017), PMLR 54, 1273–1282. https://proceedings.mlr.press/v54/mcmahan17a.html

[6] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774. https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

[7] Univora IT Group LLC. (2025). Proctoring.uz Platform Technical Documentation (Release 3.2). Tashkent: Univora IT Group LLC.

[8] OECD. (2024). OECD Principles on AI: Updated Framework for Trustworthy AI Systems (2nd ed.). OECD Publishing. https://doi.org/10.1787/oecd-ai-principles-2024

[9] O'zbekiston Respublikasi. (2019). Shaxsiy ma'lumotlar to'g'risida qonun. Toshkent: O'zbekiston Respublikasi Oliy Kengashi. https://lex.uz/

[10] VirtualAccelerate / U-Solve Incubation Programme. (2026). Certificate of Achievement: Proctoring.uz — Top 10 of 20 Projects. Tashkent: U-Enter Innovation Centre.
[11] Lancaster, T., & Cotarlan, C. (2021). Contract cheating by STEM students through a file sharing website: a Covid-19 pandemic perspective. International Journal for Educational Integrity, 17(1), Article 3. https://doi.org/10.1007/s40979-021-00070-0
[12] Swauger, S. (2021). Our bodies encoded: Algorithmic test proctoring in higher education. In J. Stommel, C. Friend, & S. Morris (Eds.), Critical Digital Pedagogy. Hybrid Pedagogy Inc.
[13] Grand View Research. (2023). Online Proctoring Market Size, Share & Trends Analysis Report 2023–2030. Grand View Research. https://www.grandviewresearch.com/industry-analysis/online-proctoring-market
[14] Nguyen, H. M., & Goto, D. (2024). On measuring the prevalence of AI cheating: Evidence from a list experiment. Assessment & Evaluation in Higher Education. https://doi.org/10.1080/02602938.2024.2417213
[15] Selwyn, N. (2019). Should Robots Replace Teachers? AI and the Future of Education. Cambridge: Polity Press.