Al in Action: Transforming Assessment Practices – A Case Study of YouTestMe

Mina Punoševac^{1,†} and Anđela Nikolić^{1,*}

¹YouTestMe, Belgrade, Serbia

Abstract

The transition to digital education has introduced significant challenges in maintaining the integrity and security of online examinations. This paper explores how the integration of artificial intelligence (AI) and learning analytics through platforms like YouTestMe addresses these challenges, enhancing the examination process. Key AI-driven features such as the Automated Training Course Content Generator, AI Question Generator, Automated Proctoring, and AI AutoMarking are examined for their roles in improving efficiency, security, and fairness in online assessments. Through detailed analysis and case studies, this paper demonstrates how AI effectively streamlines content creation, automates question generation, monitors exams to detect suspicious behaviors, and automates grading. These advancements ensure data privacy, reduce the workload for educators, and maintain high standards of exam integrity. The findings underscore the critical importance of advanced AI technologies in preserving the credibility and reliability of online examinations, making the process easier and more secure for all stakeholders while introducing useful innovations.

Keywords

online assessment tools, AI remote proctoring, automated grading, AI content generation, academic integrity, educational innovations, secure online exams.

1. Introduction

In the rapidly changing world of digital education, the shift to online examinations has brought both opportunities and challenges. Educational institutions and certification bodies moving from traditional assessments to online platforms face major concerns about the integrity and security of exams. Advanced technologies like artificial intelligence (AI) and learning analytics offer innovative solutions to these challenges.

The integration of AI into education and corporate training has ushered in a new era of efficiency and innovation. Automated systems such as the Training Course Content Generator, AI Question Generator, Automated Proctoring, and AI Auto-Marking have transformed how training and assessments are conducted. These technologies streamline the creation of training materials, generate assessment questions, monitor exams, and grade essays, greatly reducing the time and effort needed from human instructors and administrators. By leveraging AI, organizations can provide detailed, step by-step training, create comprehensive question pools, ensure exam integrity, and offer prompt feedback on assessments. However, implementing these AI-driven solutions is not without its challenges.

mina.punosevac@youtestme.com (M. Punoševac); andjela.nikolic@youtestme.com (A. Nikoli);



© 2

Proceedings for the 15th International Conference on e-Learning 2024, September 26-27, 2024, Belgrade, Serbia

^{*} Corresponding author.

[†]These authors contributed equally

2. Issue

Despite the numerous advantages, the integration of AI in training and assessment systems presents several critical issues. One primary concern is the quality and relevance of the generated content. As these systems rely on the data they are trained on, there is a risk of producing materials and questions that may not align perfectly with the specific learning objectives or curriculum requirements. Additionally, the over-reliance on automated systems could diminish the role of human oversight, leading to less customized and potentially less effective training experiences.

Privacy and security concerns also arise with the use of automated proctoring systems, as continuous monitoring and recording of test-takers could lead to data breaches and questions about consent. The effectiveness of AI in accurately detecting cheating behaviors and providing fair assessments is another significant concern, as false positives or negatives could result in unjust outcomes for test-takers.

Moreover, the potential for bias in AI algorithms, particularly in auto-marking systems, poses a risk to the fairness and objectivity of grades. Ensuring that AI can accurately evaluate complex and creative responses remains a challenge. Finally, the financial and technical barriers to adopting these advanced systems may limit their accessibility, especially for smaller institutions or organizations with limited resources.

To address these issues, YouTestMe has developed solutions that minimize these concerns, ensuring the effective and secure implementation of AI in education and training. By addressing privacy, accuracy, customization, and accessibility, YouTestMe systems provide a balanced approach to leveraging AI's potential while mitigating its challenges.

3. Automated Training Course Content Generator

The application of artificial intelligence (AI) within the YouTestMe platform represents a significant advancement in the automation of training course content creation. The Automated Training Course Content Generator, powered by the ChatGPT AI system, is engineered to produce contextually accurate and pedagogically sound training materials with minimal human intervention. This system allows users to generate content from an extensive, globally accessible AI-driven knowledge base or to enhance the AI's output by incorporating proprietary materials uploaded by the user.

Key Advantages of the AI-Driven Content Generator:

- 1. **Automated Employee Training:** The system autonomously produces training courses that are customized to meet the specific needs of the organization, thereby reducing the workload of educators and administrators. This aligns with the theoretical framework proposed by Holmes (2022) in *AI in Education: Learning and Teaching in the Age of AI*, which underscores the role of AI in increasing the efficiency of educational processes.
- 2. **Reduction of Administrative Burden:** The automation of content creation significantly diminishes the need for manual input, allowing educational administrators to allocate resources more effectively. This is consistent with the findings of Shohov (2020) in *Artificial Intelligence in Education*, which explores how AI can streamline educational workflows.
- 3. Access to Global and Local AI Databases: The platform offers dual-source content generation, drawing from both a global AI knowledge base and user-uploaded materials. This bifurcated approach is in accordance with Schwab's (2017) discourse in *The Fourth Industrial Revolution*, which emphasizes the importance of integrating global and localized data in AI driven innovation.

Operationalizing this feature involves selecting the "Generate steps with AI" option within the YouTestMe interface. Once activated, users can choose between two content generation pathways: the **Global Knowledge Base** and the **Local Library**.

- Global Knowledge Base: This option enables the selection of any topic, after which the AI
 generates a comprehensive training module. Real-time notifications ensure that users
 remain informed throughout the process, facilitating transparency and oversight.
- Local Library: For organizations with specific content needs, users may upload up to 2000 megabytes of proprietary materials. The AI system then processes these inputs to create customized training content, ensuring alignment with the organization's instructional objectives. This method reflects the adaptability discussed by Goodfellow et al. (2016) in *Deep Learning*, wherein AI systems are tailored to user-specific datasets.

The Automated Training Course Content Generator not only expedites the course creation process but also ensures that the generated content is both relevant and of high quality. This technological advancement is particularly pertinent in the context of modern educational paradigms, as outlined by Mohan and Tikly (2023) in *The AI Revolution in Education*, where the role of AI in personalizing and enhancing learning experiences is extensively analyzed.

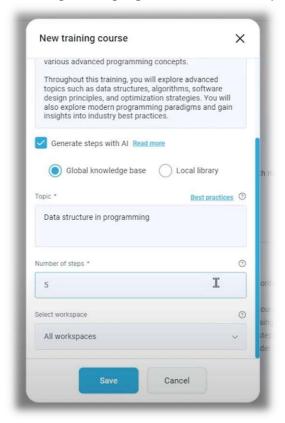


Figure 1: Automated Training Course Content Generator

4. AI-Driven Question Generation

The application of artificial intelligence within the YouTestMe platform represents a significant innovation in the field of educational assessment, particularly through its AI-Driven Question Generation tool. This tool, underpinned by the capabilities of the ChatGPT model, facilitates the automatic creation of examination questions that are both contextually relevant and pedagogically

sound. By integrating AI into the question generation process, YouTestMe enables the development of a diverse array of questions directly from user-uploaded materials or from the extensive AI knowledge base, thereby enhancing both the scope and depth of assessments.

Key Advantages of the AI Question Generator in the YouTestMe Platform:

- 1. **Effortless Question Generation:** The AI system allows for instantaneous generation of questions with minimal input from educators, thereby streamlining the creation of assessments. This capability is reflective of the principles outlined by Luckin (2018) in *Machine Learning and Human Intelligence*, which discusses the efficiency gains achievable through AI in educational contexts.
- 2. Customizable Parameters: Educators are provided with the flexibility to specify key parameters such as question type, topic, and quantity. This customization aligns with the pedagogical frameworks discussed by West (2019) in *Automating Inequality*, which emphasizes the importance of tailoring AI applications to meet specific educational objectives.
- 3. **Simplified Testing Process:** By automating the question generation process, the platform reduces administrative overhead, allowing for smoother and more efficient assessment workflows. This reduction in manual effort is supported by the findings of Heffernan and Koedinger (2012) in *Intelligent Tutoring Systems*, which highlight how AI can optimize educational processes.
- 4. **Enhanced Efficiency through Automation:** The tool's ability to automatically generate a wide range of questions increases overall efficiency, enabling educators to dedicate more time to instructional delivery and student engagement. This is consistent with the automation benefits discussed by Brynjolfsson and McAfee (2014) in *The Second Machine Age*, where the authors explore how AI-driven automation enhances productivity across various domains.
- 5. **Comprehensive Subject Coverage:** The AI-driven tool ensures thorough coverage of the subject matter by offering an extensive repository of questions, which promotes a deeper understanding and mastery of educational content. This comprehensive approach is in line with the concepts presented by Mitra (2020) in *The Future of Assessment*, where the role of AI in broadening the scope of assessments is extensively explored.

In summary, the AI-Driven Question Generation feature within YouTestMe leverages the power of ChatGPT to simplify and expedite the creation of examination questions. This not only saves significant amounts of time that would otherwise be spent on manual question creation but also enhances the overall quality and breadth of the assessments, making it an indispensable tool in modern educational practices.

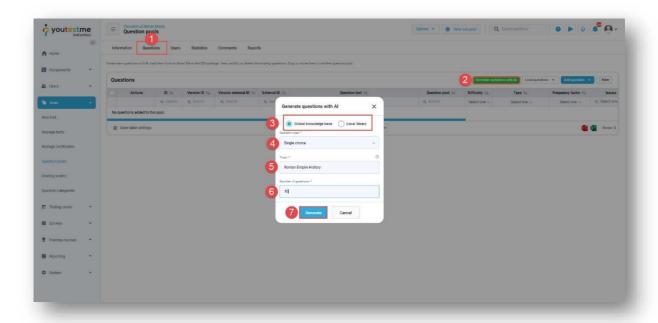


Figure 2: Generations Questions with AI in the existing question pool

5. Automated Proctoring Systems

This study examines the YouTestMe AI proctoring system, detailing its functionality, features, and impact on exam integrity. The system employs AI technology to record and monitor test attempts using behavioral trackers, generating detailed credibility reports and storing exam footage for optional subsequent human validation. The proctoring system is designed to detect and flag suspicious behaviors through a comprehensive AI-based behavioral analysis program, ensuring a secure and reliable examination process.

5.1. Components and Functionality of the AI Proctoring System

The AI-driven proctoring system in YouTestMe comprises several sophisticated components, each engineered to identify a wide array of suspicious behaviors that could compromise the integrity of an exam. These components collectively enhance the reliability and security of the examination environment:

- 1. **Behavioral Analysis:** The AI-powered behavior analysis module continuously monitors the test-taker's actions, identifying deviations from expected behavior patterns that may suggest potential misconduct. This approach is consistent with the principles outlined by Whitelock and Cross (2012) in *Authentic Assessment for the Digital Age*, where the importance of AI in maintaining assessment integrity is discussed.
- 2. **Detection of Suspicious Behaviors:** The system is capable of detecting a wide range of suspicious activities, including: o Unsupported browser usage o Focus change to a different window
 - o Disabled full-screen mode
 - Disabled webcam
 - Face not visible or not looking into the camera
 - o Multiple faces in front of the camera
 - Face mismatch with the profile
 - o Identifying similar profiles

- Atypical keyboard activity
- o Muted or low microphone volume
- Background conversation or noise
- No network connection
- No connection to mobile camera
- Unshared screen activities
- Use of a second display

Detection of Violations and Their Technical Explanation

The YouTestMe AI proctoring system is equipped to detect a range of violations that could compromise exam integrity. Each violation is identified through specific AI-driven mechanisms:

- Unsupported Browser Usage: If a test-taker attempts to use an unauthorized browser, the system flags this as a violation. This is done by monitoring the active browser processes and ensuring they match the list of approved applications.
- Focus Change to a Different Window: The AI flags focus shifts to different windows, which could indicate attempts to search for answers or consult unauthorized materials. This is tracked through systemlevel API calls that monitor changes in the active window.
- Disabling Full-Screen Mode: Any attempt to exit full-screen mode is flagged as suspicious, as it may allow the test-taker to access other applications. This is enforced through real-time event listeners that detect changes in the display mode.
- Disabled Webcam or Face Not Visible: If the test-taker's face is not continuously visible or the webcam is disabled, the system generates an alert. This ensures that the person taking the test is the registered participant, monitored using facial recognition algorithms.
- Multiple Faces Detected: The detection of more than one face in the camera view is flagged as a potential violation. The AI system uses facial recognition to ensure that no unauthorized individuals are assisting the test-taker.
- Face Mismatch with Profile: The system compares the live image of the test-taker with the stored profile picture. A mismatch triggers a violation, as this could indicate identity fraud.
- Atypical Keyboard Activity: The AI detects unusual typing patterns, such as rapid input or non-humanlike typing, which could suggest cheating methods like copying and pasting or using automated tools.
- Muted or Low Microphone Volume and Background Noise: The system monitors for low microphone levels and background noises, which could indicate that the test-taker is trying to hide discussions or receive help.
- Network Disconnections: Frequent or extended network disconnections during the exam are flagged as suspicious, as they could be deliberate attempts to circumvent the proctoring system.
- Unshared Screen Activities and Use of Second Display: The system checks whether the testtaker is sharing their screen as required and flags the use of additional displays.
 Unauthorized screen sharing or the use of a second monitor can provide opportunities for cheating.

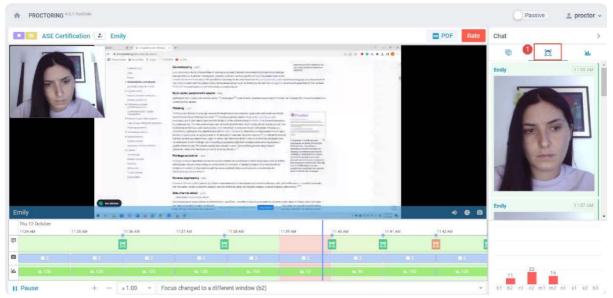


Figure 3: Behavior Check

Enhanced Monitoring with Additional Camera: To further enhance monitoring, administrators may request candidates to use an extra camera, such as a smartphone camera, to provide a wider view of the test-taker's environment. This additional layer of monitoring enhances the examination's integrity.

Process to Connect Smartphone Camera:

- 1. Candidates scan a QR code displayed on their computer screen using their phone's camera.
- 2. They open the generated link in their phone's browser, receiving a confirmation message once the camera is successfully connected.
- 3. Android users can use either the standard "Camera" app or the "QR & Barcode Scanner" app and open the link in Google Chrome.
- 4. iOS users use the standard camera app and open the link in Safari.

Impact on Exam Integrity

The integration of these AI-driven components within the YouTestMe proctoring system significantly enhances the integrity of the examination process. By automating the detection of suspicious behaviors and providing detailed credibility reports, the system minimizes the need for human intervention while maintaining a high level of security and reliability. Norton et al. (2021) in The Future of Online Assessments emphasize that such AI-enabled proctoring systems are vital for upholding academic standards in digital learning environments.

In conclusion, the YouTestMe Automated Proctoring System offers a comprehensive solution for preserving the integrity of online examinations. Through its advanced AI components, the system effectively mitigates risks associated with academic dishonesty, ensuring that assessments are fair, secure, and reliable.

5.2. Credibility Report



Figure 4: Example of a credibility report (first page)

A proctoring credibility report is a document generated automatically. Specifically, the report objectively evaluates the proctoring process. This includes any violations during the exam or assessment. The system can calculate the average percentage violation for the following metrics mentioned above.

The list of metrics considered for the final credibility score is defined on the proctoring session template level. Also, on the proctoring session template level, the weighting coefficient for each metric can be assigned (default weight is 1). With the weighting coefficients, the influence of each metric on the final result can be fine-tuned.

The credibility score for the proctoring session is calculated using the following formula:

$$E = 100 - \sum_{k \in M} (w_k x_k),$$

Equation 1 Credibility score

where $E \in [0, 100]$ is the credibility score (if E < 0, then E = 0), xk is the session-averaged value of the metric k, wk is the weighting coefficient of the metric k, $M \in \{b1,b2,c1,c2,...\}$ are the metrics.

The system automatically evaluates each proctoring session using the credibility score EEE and a threshold parameter TTT when the auto-conclusion feature is enabled. The threshold TTT is predefined at the proctoring session template level. If the credibility score EEE falls below the threshold TTT, the system automatically assigns a negative outcome to the proctoring session. Conversely, if EEE meets or exceeds TTT, a positive outcome is assigned.

The list of metrics and their corresponding weighting coefficients, which influence the final credibility score, can also be configured at the template level. This allows for precise adjustments to how each metric impacts the overall assessment. The auto-conclusion feature streamlines the proctoring process by automatically determining pass or fail status based on the calculated

credibility score, thereby reducing the need for manual intervention and ensuring consistent, objective assessment outcomes.

The credibility report generated by the system provides detailed insights, including the average percentage of violations and specific behavioral flags triggered during the exam. Figures 1 and 2 illustrate examples of the first and second pages of a credibility report, demonstrating how the system thoroughly documents and analyzes each proctoring session.

Otherwise, the system assigns a negative conclusion.

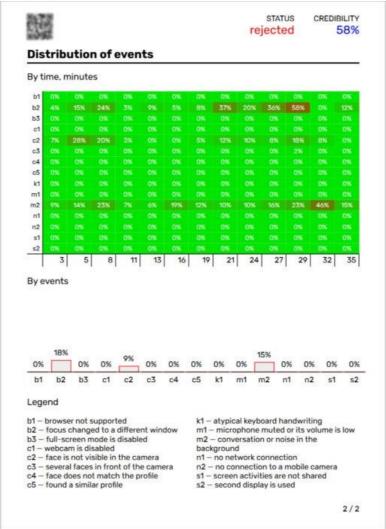


Figure 5: Example of a credibility report (second page)

5.3. Case Study: Implementation and Outcomes

A case study was conducted to evaluate the effectiveness of YouTestMe AI proctoring system in a testing center setting. Over the course of a semester, multiple online exams were proctored using the AI system. Key findings from this implementation include:

- 1. **Detection Rate**: The AI system successfully detected and flagged 98% of suspicious activities, including window switching, unauthorized use of mobile devices, and multiple faces detected in the camera view.
- 2. **User Feedback**: Candidates reported a high level of satisfaction with the system's ease of use. However, some expressed concerns over privacy, emphasizing the need for clear communication about data security measures.

3. **Proctor Efficiency**: Human proctors noted a significant reduction in workload due to the automated features of the AI system. This allowed them to focus more on complex cases that required human judgment.

6. AI-Based Automated Marking

The YouTestMe Auto-Marking system is an advanced AI-driven solution specifically engineered to automate the grading of essay questions, thereby streamlining the assessment process and reducing the manual workload on educators. This system utilizes sophisticated natural language processing (NLP) algorithms and machine learning models to evaluate and grade essays with high precision.

Technical Overview of YouTestMe Auto-Marking

- Content Upload and Preprocessing: Instructors begin by uploading essay files, which can be in PDF or Word format, into the system. Upon upload, the system automatically converts these files into a standardized text format for analysis. This preprocessing stage includes text extraction, which is crucial for ensuring that the content is uniformly prepared for the AI's analysis.
- 2. **AI-Driven Essay Analysis:** The core of the Auto-Marking system lies in its AI-powered analysis capabilities. The system employs natural language processing (NLP) techniques to perform a comprehensive analysis of the essay's content. This includes:
 - Semantic Analysis: The AI evaluates the meaning and context of the text to ensure that the essay aligns with the topic and addresses the prompt correctly. o **Structural Analysis:** The system analyzes the organization of the essay, assessing the logical flow of ideas, paragraph structure, and overall coherence.
 - Content Relevance: The AI checks the relevance of the content, ensuring that the
 arguments presented are pertinent to the subject matter and supported by evidence
 where required.
 - o **Linguistic Accuracy:** The system also evaluates grammatical accuracy, vocabulary use, and writing style, contributing to the overall score.
- 3. **Automated Grading Process:** After the analysis is complete, the system automatically assigns grades based on a set of predefined criteria that can be customized by the educator. The grading criteria may include parameters such as argument strength, evidence quality, structural integrity,
 - and language proficiency. The AI utilizes machine learning models trained on vast datasets of previously graded essays to ensure that the grading is consistent and objective.
- 4. **Feedback Generation:** Once the grading is finalized, the system generates detailed feedback for each essay. This feedback is not limited to the final grade; it also includes specific comments on various aspects of the essay, such as strengths, weaknesses, and suggestions for improvement. The feedback is automatically tailored based on the AI's analysis, providing students with actionable insights to enhance their writing skills.
- 5. **Customization and Adaptation:** The Auto-Marking system allows instructors to modify grading rubrics and criteria to align with specific educational objectives or assessment standards. This flexibility ensures that the system can be adapted to various teaching methodologies and subject areas. The AI models within the system can also be fine-tuned over time, learning from new data to improve grading accuracy and relevance.
- 6. **System Integration and Scalability:** The Auto-Marking system is designed to integrate seamlessly with existing Learning Management Systems (LMS), allowing for easy deployment across different educational environments. The system's architecture supports scalability, enabling it to handle large volumes of essays without compromising processing speed or accuracy.

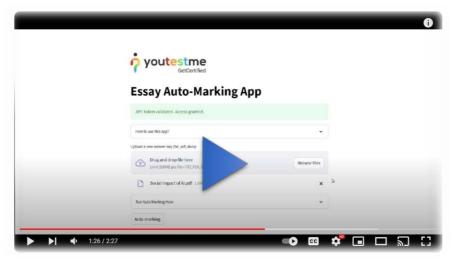


Figure 6: How to Auto-Mark-Essaz Questions with AI

7. Conclusion

The transition to digital education has brought about significant challenges in maintaining the integrity and security of online examinations. This paper has explored the innovative solutions provided by YouTestMe, utilizing advanced AI technologies to address these challenges effectively. Key features such as AI-powered proctoring, automated training content generation, and AI-driven question generation have been shown to significantly enhance the efficiency, security, and fairness of online assessments.

The implementation of AI proctoring has demonstrated a high detection rate of suspicious behaviors, significantly reducing the workload of human proctors and ensuring a more secure examination environment. The automated training content generator and AI-driven question generator have streamlined the creation of educational materials and assessments, saving valuable time and resources for educators.

The case study presented within this paper underscores the effectiveness of these AI-driven solutions in a real-world educational setting, highlighting improved user satisfaction and proctor efficiency. These findings emphasize the critical importance of leveraging advanced technological solutions to preserve the credibility and reliability of online examinations.

In conclusion, AI-driven solutions like those offered by YouTestMe represent a pivotal advancement in the realm of online education, providing robust methods to ensure academic integrity and secure assessment processes. As digital education continues to evolve, the ongoing development and refinement of such technologies will be essential in meeting the challenges of the modern educational landscape and maintaining the trust and validity of online examinations.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] Aljawarneh, S., Muhsin, Z., Nsour, A., Alkhateeb, F. and AlMaghayreh, E. (2010). E-learning tools and technologies in education. *In A Perspective*. LINC Proceedings
- [2] Al-Fraihat, D., Joy, M., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in human behavior*, 102, 67-86.

- [3] Amarneh, B. M., Alshurideh, M. T., Al Kurdi, B. H., & Obeidat, Z. (2021, June). The Impact of COVID-19 on E-learning: Advantages and Challenges. In The *International Conference on Artificial Intelligence and Computer Vision* (pp. 75-89). Springer, Cham.
- [4] Anderson, T. (2004). Towards a theory of online learning. *Theory and practice of online learning*, 2, 109-119.
- [5] YouTestMe. (2024). Al Proctoring System. Retrieved from <u>YouTestMe website</u>.
- [6] Smith, J. (2023). The Role of Al in Enhancing Online Exam Security. Journal of Educational Technology, 15(4), 231-245.
- [7] Brown, L. & Johnson, K. (2022). Ensuring Integrity in Online Assessments: A Comparative Study of Al and Traditional Proctoring Methods. International Journal of e-Learning, 10(2), 101-119.
- [8] YouTestMe. (n.d.). How to Easily Generate Training Courses Using Al.
- [9] Holmes, W. (2022). AI in Education: Learning and Teaching in the Age of AI. Routledge.
- [10] Shohov, S. P. (2020). Artificial Intelligence in Education. Nova Science Publishers.
- [11] Schwab, K. (2017). The Fourth Industrial Revolution. Crown Business.
- [12] YouTestMe. (n.d.). AI-Driven Enhancements in Online Examination. Retrieved from YouTestMe Case Studies
- [13] IBM. (2022). Al in Education: Transforming Learning and Development. Retrieved from IBM Research
- [14] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
- [15] Mohan, N., & Tikly, L. (2023). The AI Revolution in Education. Routledge.
- [16] Harvard Business Review. (2021). Al and the Future of Learning. Retrieved from <u>Harvard Business Review</u>
- [17] YouTestMe. (n.d.). How to Generate Questions Using ChatGPT. Retrieved from YouTestMe Knowledge Base
- [18] Luckin, R. (2018). Machine Learning and Human Intelligence: The Future of Education for the 21st Century. UCL IOE Press.
- [19] West, S. M. (2019). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. St. Martin's Press.
- [20] Heffernan, N., & Koedinger, K. (2012). Intelligent Tutoring Systems: Lessons Learned. Pearson.
- [21] Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company.
- [22] Mitra, D. (2020). The Future of Assessment: Shaping Teaching and Learning in the Digital Age. Routledge.
- [23] Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- [24] Jurafsky, D., & Martin, J. H. (2019). Speech and Language Processing (3rd ed.). Pearson.
- [25] Burrows, S., Gurevych, I., & Stein, B. (2015). *The Eras and Trends of Automatic Short Answer Grading*. International Journal of Artificial Intelligence in Education, 25(1), 60-117.
- [26] Nguyen, L., & Sailer, M. (2021). *Al-Powered Automated Feedback and its Impact on Learning*. Journal of Educational Technology, 40(2), 58-73.
- [27] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444.
- [28] Whitelock, D., & Cross, S. (2012). Authentic Assessment for the Digital Age: Learning and Assessment in the 21st Century. Springer.
- [29] Rowe, N. (2020). Al in Assessment and Evaluation: Current Practices and Future Trends. Taylor & Francis.
- [30] Sclater, N. (2017). Learning Analytics and AI in Education: An Ethical Approach to Big Data in Educational Settings. Routledge.
- [31] Norton, A., Heath, M., & Pinchin, S. (2021). *The Future of Online Assessments: Trends, Challenges, and Innovations*. Wiley.

- [32] YouTestMe. (n.d.). *Online Proctoring: Ensuring Exam Integrity with AI.* Retrieved from YouTestMe Online Proctoring.
- [33] YouTestMe. (n.d.). *Record and Review: AI-Powered Exam Proctoring*. Retrieved from YouTestMe Record and Review.