Funding was provided by the Research Council of Lithuania (LMTLT), contract No. S-ITP-25-14.

Artificial Intelligence and Multimodal Data Fusion System for Assessing and Detecting Fraud in Applicants` Videos

Diana Kalibatiene*, Algirdas Laukaitis, Kęstutis Normantas, Julius Jancevičius*, Mindaugas Jankauskas, Dovilė Jodenytė

Department of Information Systems, Faculty of Fundamental Sciences, Vilnius Gediminas Technical University, * diana.kalibatiene@vilniustech.lt, julius.jancevicius@vilniustech.lt

ABSTRACT

- **Topic/Context:** The process of applying to international universities requires submitting multiple documents and a video interview recording.
- The applicant's data is collected in diverse and multimodal formats: scanned images, text documents (Word, PDF), and video recordings of interviews.
- **Current Challenges:**
  - Processing this multimodal data – a time-consuming and human resource-intensive process.
  - Applicant selection decisions may be subjective and do not reveal the real level of his/her preparation, as the process is often carried out by only one person.
- **Research Aim –** to enhance the applicant assessment process and to minimize the assessment bias by developing an AI-based assessment system – FAIR-VID.

MAIN CONCEPTS

- **Asynchronous Video Interview** (AVI) - candidates respond to computer-delivered interview questions, recorded via their device's camera without any interaction with interviewers, and these responses are then uploaded to a central platform for review by employers [4].

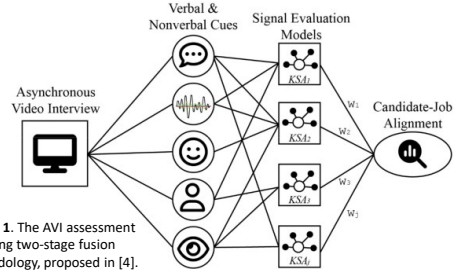

Figure 1. The AVI assessment using two-stage fusion methodology, proposed in [4].

Note: Figure 1 presents the process of extracting verbal and nonverbal cues from AVI recordings and then using these cues as input into separate models to automatically assess job-relevant knowledge, skills, and abilities. The Signals Evaluation Models implemented and their weighted values will vary based on job-analyses performed by organizations.

ACKNOWLEDGEMENT

These results are part of the project "Artificial intelligence and multimodal data fusion system for assessing and detecting fraud in applicants` videos" (FAIR-VID). This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-25-14.

PROPOSED AI-BASED ADMISSION WORKFLOW

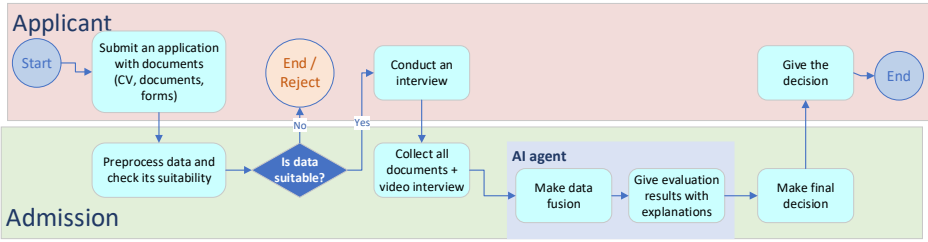

Figure 2. The AI-based admission workflow

- The FAIR-VID project addresses the challenge of ensuring data privacy and security while enabling AI-driven analysis of sensitive admissions data. The mind map (Figure 3) organizes the main implementation paths for privacy-preserving AI, each linked to specific technologies and frameworks that support secure, compliant, and effective workflows.

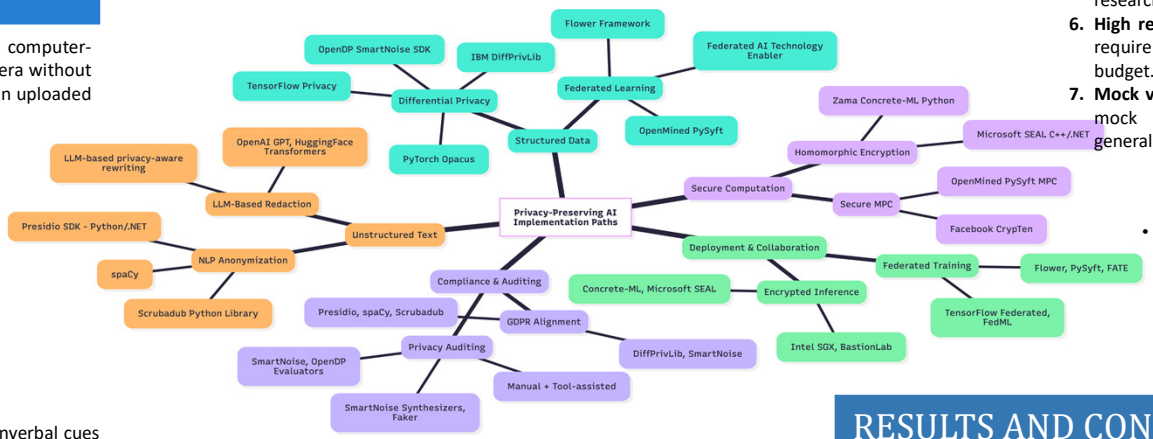

Figure 3. The mind map of the main implementation paths for privacy-preserving AI

LITERATURE

1. Hickman, L.; Saef, R.; Ng, V.; Woo, S.E.; Tay, L.; Bosch, N. Developing and Evaluating Language-Based Machine Learning Algorithms for Inferring Applicant Personality in Video Interviews. *Human Resource Management Journal* 2024, 34, 255–274, doi:10.1111/1748-8583.12356.

2. Todsam, S.; Kerkar, S.; Valte, A.; Masal, K. Mock Interview Platform for Interview Performance Analysis Using Machine Learning.; Institute of Electrical and Electronics Engineers (IEEE), September 24 2025; pp. 1201–1206.

3. Uppalapati, P. J., Dabbiru, M., & Kasukurthi, V. R. (2025). AI-driven mock interview assessment: leveraging generative language models for automated evaluation. *International Journal of Machine Learning and Cybernetics*, 1-23.

4. Pentland, S. J., Wang, X., & Twyman, N. W. (2025). Better job application systems: Objectively assessing measures of job performance from asynchronous video interviews. *Information & Management*, 62(2), 104077.

ISSUES TO BE SOLVED

1. **Transcript-only evaluation** – Current automated interview grading ignores non-verbal cues (facial expressions, body language) [1].
2. **Cultural & linguistic barriers** – Limited research on non-native speakers, accents, demographic bias, and multilingual contexts [2].
3. **Dataset limitations** – Lack of sophisticated open-ended question datasets with technical domain content [3].
4. **Performance focus only** – Studies emphasize algorithm performance over actual job performance correlation.
5. **Security gaps** – Insufficient attention to anonymization and security due to varying research focuses.
6. **High resource costs** – LLMs for interview grading require significant computational resources and budget.
7. **Mock vs. real interviews** – Analysed methods use mock interviews, limiting real-world generalizability.

- This structured approach (Figure 3) ensures that FAIR-VID can leverage advanced AI for fairness and utility while rigorously protecting personal data, complying with legal and ethical standards, and maintaining trust among stakeholders.

RESULTS AND CONCLUSIONS

1. Multimodal AI fusion of video interviews and documents shows strong potential for automating university admissions while addressing scalability, objectivity, and bias reduction.
2. Audio, visual, and verbal features from AVIs demonstrate substantial predictive validity ($R^2 = 0.44$) and personality traits explain significant interview variance ($R^2 = 0.32$), with verbal features showing strongest explanatory power.
3. Future research should validate findings in high-stakes contexts, integrate objective performance criteria, address demographic bias and multilingual support, and incorporate advanced techniques like contemporary LLMs to enhance accuracy while maintaining ethical standards.