School of Science, Engineering & Technology

RMIT UNIVERSITY — School of Science, Engineering & Technology

AviAi: AI-Powered Aviation Cognitive Interview

Team NEUROWINX - SGS

- Nguyen Tran Ha Anh - s3938490
- Truong Hong Van - s3957034
- Pham Le Quynh Anh - s3927427
- Bui Tuan Anh - s3970375
- Tran Ly The Quang - s3878707

Academic Supervisor

- Dr. Arthur Tang
- Dr. Marco D'Alessandro
- Dr. Thanh Pham
- Mr. Tom Huynh

Project Background & Motivation

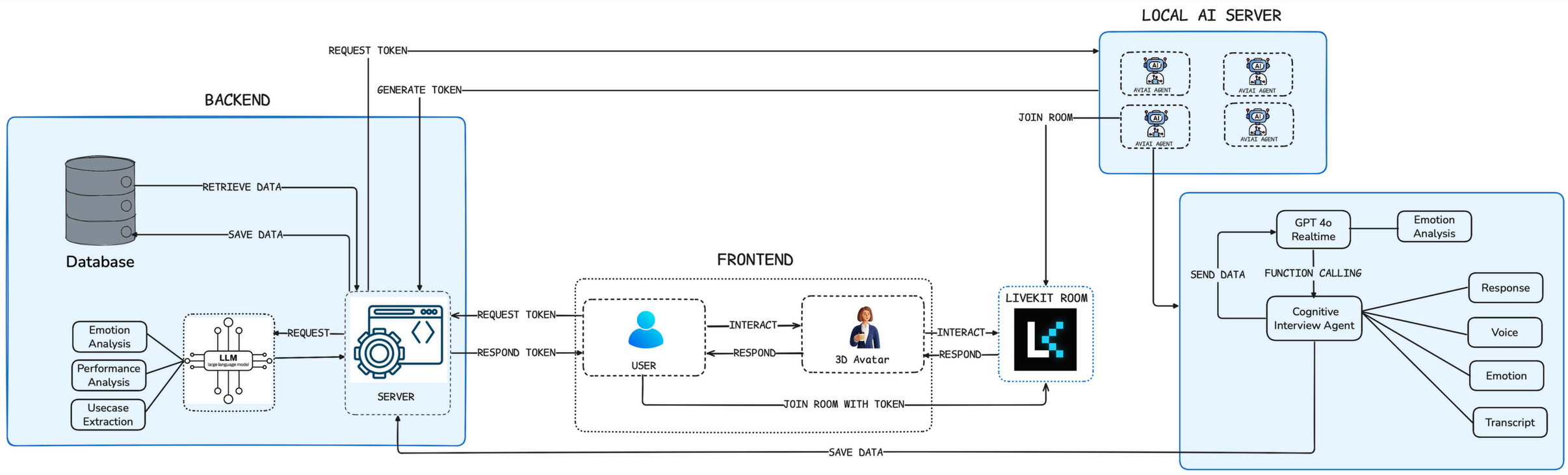
In high-stakes fields like aviation accident investigation, the Cognitive Interview (CI) is a crucial technique for obtaining accurate eyewitness accounts, overcoming the natural fallibility of human memory. However, mastering this skill requires extensive practice, and current training methods are inadequate. Trainees typically rely on traditional role-playing, which is logistically challenging, unscalable, and poses a significant ethical risk of vicarious traumatization to participants.

This project is motivated by this critical training gap. We are developing an AI-driven avatar platform to provide a safe, accessible, and realistic simulation environment. This will allow future investigators to effectively build their skills and confidence without the practical and psychological drawbacks of traditional approaches, ensuring they are better prepared for real-world challenges.

Objectives

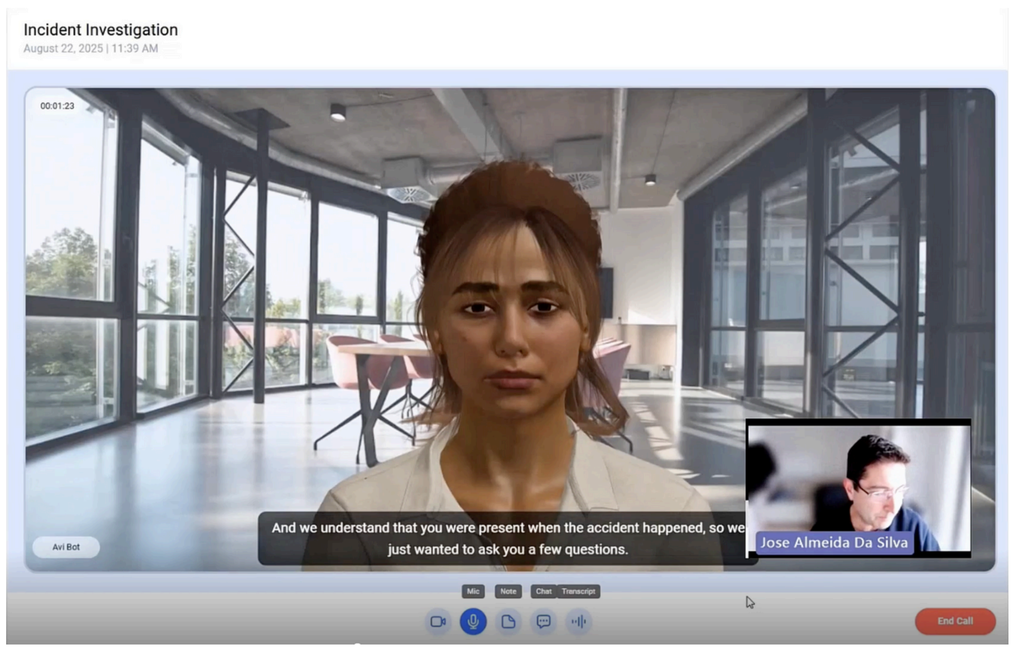
- Develop an AI-driven cognitive interview training platform featuring a realistic, emotionally expressive 3D avatar powered by LLMs, speech recognition, and voice synthesis to simulate authentic eyewitness interviews in aviation accident investigations.
- Enhance student learning and engagement by providing an immersive, scalable, and psychologically safer alternative to traditional role-play training, improving skill acquisition while reducing ethical and emotional risks.
- Create a consistent, accessible, and adaptable training tool that can extend beyond aviation to fields such as law enforcement, forensic psychology, and emergency response

Methodology



1. **Connect & Join Session**
   - User use a session's token and connect to a LiveKit room
   - The AI server dispatches a dedicated agent to join the session
2. **Real-time interaction Loop**
   - User's audio is streamed to GPT-4o for instant transcript and emotion analysis
   - Our LLM Agent use the scenario's context and converstaion history to reason and generate a response
   - The Agent output a JSON object with speech, vocal tone, and facial expressions for the 3D avatar
3. **Save & Analysis**
   - When user end the interview, the full conversation history is sent to our backend and saved in the database for further analysis

Experiment & Results



Our experiments were designed to validate AviAI's technical performance and user experience across iterative development versions. Using a standardized interview scenario, we measured key metrics including response latency, adherence to Cognitive Interview (CI) principles, and user-perceived realism. The results demonstrate significant improvements with each version, validating our architectural refinements. The transition to more sophisticated LLM prompting and optimized LiveKit handling directly correlated with lower latency, more accurate CI technique application, and a more immersive, natural-feeling simulation for the user.

No	Metric	Version 1.0 (Baseline)	Version 2.0 (Prompt Engineering)	Version 3.0 (Optimized LiveKit & Prompt)
1	Avg. Response Latency	7.8s	4.8s	1.2-1.8s
2	CI Technique Adherence	40%	75%	87%
3	User Realism Rating (1-5)	2.8 / 5	3.3 / 5	3.9 / 5
4	Connection Stability	Low	Moderate	High

Conclusion & Future work

In conclusion, this project successfully developed AviAI, a functional AI-powered platform designed to address the critical gaps in Cognitive Interview (CI) training. By creating an interactive and psychologically safe simulation, we delivered on our objective to provide an effective, modern learning tool that bridges the gap between theory and practice. AviAI transforms a high-risk, logistically complex training process into a scalable, on-demand experience that enhances skill acquisition while prioritizing the well-being of trainees. The iterative improvements to our LLM prompting and LiveKit integration have proven to increase performance, realism, and adherence to CI principles, validating the system's core architecture and its potential as a powerful educational asset.

To build upon the success of this project, we have identified three key areas for future development:

- Deeper Interviewer Analysis: Integrate eye-tracking and body language detection to provide more holistic feedback on the trainee's non-verbal cues.
- Enhanced Avatar Realism: Improve the avatar's visual fidelity and significantly expand its range of emotional expressions for a more immersive experience.
- Advanced Performance Dashboard: Develop a detailed feedback system to provide students with actionable analysis of their interview performance, helping them track their progress.