**Autonomous Interview Screening System**

**Abstract**

This report outlines my contributions to an autonomous interview screening system developed for my master’s dissertation. The project aimed to create an AI-driven system capable of conducting realistic interview simulations and providing comprehensive post-interview analysis. The group consists of Niranjan Mahtani Kewalramani (member), Marek Grzes (Supervisor) and me. The structure and implementation of the project was discussed and decided by all three of us. We had regular meetings with Marek every week and frequent meetings amongst ourselves both in person and on discord to make progress.

**1. Introduction**

The system comprises two main phases: "Real-Time processing" and "Post Processing" (see Appendix A, Figure 1). The former focuses on real-time candidate interaction, while the latter analyses the interview data to provide a comprehensive evaluation with detailed feedback.

**2. Speech-to-Text Conversion**

To achieve accurate and cost-effective transcription, I implemented an open source locally hosted version of OpenAI's Whisper [1]. This approach provided free and reliable multi-lingual transcription capabilities, with the "medium" model proving sufficiently accurate for our purposes.

**3. User Interface Design**

Using the Tkinter library [2], I designed a multi-screen interface (see Appendix A, Figures 2 and 3) featuring input fields for job details, adjustable parameters for the speech-to-text engine and the initial LLM(Claude 3.5 Sonnet), and a dedicated interview screen with a Speak/Stop button for the candidate to record their speech.

**4. Answer Accuracy Verification**

I developed an accuracy verification system using a Retrieval Augmented Generation (RAG) pipeline with the LangChain framework [3]. This system breaks down candidate answers into search queries, gathers relevant information through web scraping, and stores it in a vector database (Chroma) with OpenAI embeddings. It then retrieves the relevant parts for verification and adds it to the LLM's context window for assessment [4, 5].

**5. Sentiment Analysis**

After exploring custom sentiment analysis models with biased datasets, I integrated the Hume AI platform [6] for audio sentiment analysis. This provides a rich set of emotional and sentiment-related metrics, offering deeper insights into candidate responses (see Appendix B for sample metrics).

**6. Additional Contributions and System Deployment**

I implemented robust exception handling, created comprehensive documentation, collaborated on system testing and evaluation strategies, and contributed to deployment planning, including dependency management and executable packaging options.

**7. Conclusion**

Our work on this project has been instrumental in developing a functional and insightful interview screening system. By addressing practical challenges with innovative solutions, this project showcases the potential of AI in recruitment and opens avenues for further exploration in this field.

**Bibliography**

[1] Openai (no date) *Openai/whisper: Robust speech recognition via large-scale weak supervision*, *GitHub*. Available at: https://github.com/openai/whisper (Accessed: 16 August 2024).

[2] *Tkinter - Python interface to TCL/TK* (no date) *Python documentation*. Available at: https://docs.python.org/3/library/tkinter.html (Accessed: 16 August 2024).

[3] *Introduction* (no date) *🦜️🔗 LangChain*. Available at: https://python.langchain.com/v0.2/docs/introduction/ (Accessed: 16 August 2024).

[4] Pi, W. (2024) *Efficient information retrieval and response generation with retrieval-augmented generation (RAG)* [Preprint]. doi:10.59350/q2pq3-0fv85.

[6] *Welcome to Hume Ai* (no date) *Hume API*. Available at: https://dev.hume.ai/intro (Accessed: 16 August 2024).

[7] Livingstone, S.R. (2019) *Ravdess emotional speech audio*, *Kaggle*. Available at: https://www.kaggle.com/datasets/uwrfkaggler/ravdess-emotional-speech-audio (Accessed: 16 August 2024).

**Appendix A:**

**A diagram of a diagram

Description automatically generated**

Figure 1. System workflow

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**Figure 3. Job Details and Settings Screen**A screenshot of a computer

Description automatically generated**

Figure 4. Control Panel for Stop

Figure 3. Control Panel for Speak

**Appendix B:**

1. **Hume Analysis Metrics**:

Emotion Scores:  
Admiration: 0.029698949307203293  
Adoration: 0.04302431270480156  
Aesthetic Appreciation: 0.01405085064470768  
Amusement: 0.046523310244083405  
Anger: 0.004918431863188744  
Annoyance: 0.003085097298026085  
Anxiety: 0.0031014648266136646  
Awe: 0.04468310624361038  
Awkwardness: 0.0035788188688457012  
Boredom: 0.000661490426864475  
Calmness: 0.03959153592586517  
Concentration: 0.0022455144207924604  
Confusion: 0.004444075282663107  
Contemplation: 0.010204021818935871  
Contempt: 0.021374398842453957  
Contentment: 0.2674786150455475  
Craving: 0.002267987234517932  
Desire: 0.0012277523055672646  
Determination: 0.009590454399585724  
Disappointment: 0.0007131476886570454  
Disapproval: 0.0014754937728866935  
Disgust: 0.0011475816136226058  
Distress: 0.002752037486061454  
Doubt: 0.000799593806732446  
Ecstasy: 0.34356117248535156  
Embarrassment: 0.000743867305573076  
Empathic Pain: 0.002465164987370372  
Enthusiasm: 0.3080814778804779  
Entrancement: 0.019871320575475693  
Envy: 0.0008098236285150051  
Excitement: 0.3121441602706909  
Fear: 0.0009849846828728914  
Gratitude: 0.1208457499742508  
Guilt: 0.0005904586869291961  
Horror: 0.0005830202717334032  
Interest: 0.01758081093430519  
Joy: 0.8937432765960693  
Love: 0.06204117834568024  
Nostalgia: 0.0017802549991756678  
Pain: 0.001951383426785469  
Pride: 0.06803938001394272  
Realization: 0.026851702481508255  
Relief: 0.06699783354997635  
Romance: 0.005580543074756861  
Sadness: 0.001141536282375455  
Sarcasm: 0.010973036289215088  
Satisfaction: 0.2749807834625244  
Shame: 0.0010663117282092571  
Surprise (negative): 0.002550811041146517  
Surprise (positive): 0.09627598524093628  
Sympathy: 0.002557111205533147  
Tiredness: 0.0013367693172767758  
Triumph: 0.17714570462703705  
Sentimen Scores:  
1: 0.0010329830693081021  
2: 0.0010178626980632544  
3: 0.0013630579924210906  
4: 0.0026854900643229485  
5: 0.014838673174381256  
6: 0.01902943290770054  
7: 0.05759797617793083  
8: 0.1589507907629013  
9: 0.7589902877807617  
Toxicity Scores:  
identity\_hate: 0.003196313977241516  
insult: 0.0029195230454206467  
obscene: 0.0017942077247425914  
severe\_toxic: 0.0021684677340090275  
threat: 0.0032580411061644554  
toxic: 0.006401388440281153