AI-Assisted Air Traffic Conflict Resolver – LLM and AI/DS Track

GITHUB: https://github.com/JaiPatel31/ATC-Ai-HumanLoop

Jai Patel, Nico Weber

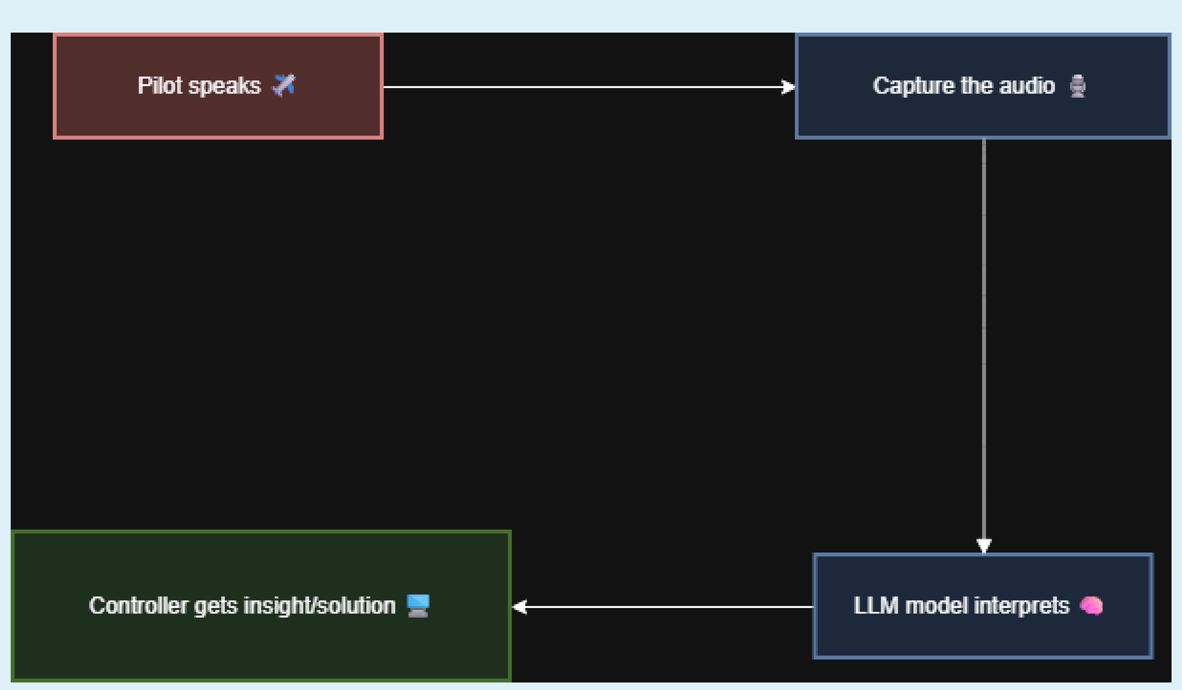
Computing, Analytics, and Mathematics, School of Science and Engineering University of Missouri – Kansas City

Introduction

Air traffic control (ATC) continues to be a major issue that plagues aviation. It is believed that there have been at least 135 accidents or incidents dating back to 2010 in which human error played a part. These incidents can come down to things like simple mistakes and lapses in judgement. The FAA has already begun testing on AI tools to help reduce the amount of runway incidents, showing growing interest in AI. By integrating AI alongside human ATC, we should be able to cut down significantly on the number of accidents and incidents that are related to human error.

Objectives

1. **Develop a voice-driven AI assistant** capable of interpreting real-time air-traffic controller and pilot communications using speech-to-text and natural language parsing.
2. **Transform unstructured radio transmissions into structured data** by extracting key flight parameters such as callsign, heading, altitude, and command.
3. **Detect and resolve potential air-traffic conflicts automatically** through a rule-based or data-driven reasoning engine that evaluates aircraft states.
4. **Generate realistic controller responses** using text-to-speech (TTS) to simulate a complete two-way communication loop.
5. **Visualize aircraft and interactions** in an intuitive React dashboard that displays transcripts, parsed data, conflict alerts, and AI-generated responses.



Methods

1. Audio Input & Preprocessing

Pilot and controller transmissions are recorded or uploaded as .wav files.

Audio is normalized and segmented into clips for transcription.

The backend (FastAPI) manages uploads and preprocessing in real time.

2. Speech-to-Text (STT)

Implemented with **Whisper Large-V3 (fine-tuned for ATC radio communications)**.

Converts noisy ATC audio into structured text.

Falls back to manual text entry when GPU or model unavailable.

3. Intent Parsing (NLP Layer)

Custom **ATC Parser** normalizes transcripts into structured data:

- **Callsign** (e.g., CSA025)
- **Command** (e.g., descend, turn, maintain)
- **Flight Level / Altitude**
- **Heading / Direction**
- **Speaker Role** (pilot or controller)

Heuristics and token mapping capture ICAO phraseology.

4. Conflict Detection Engine

Parsed aircraft data is stored in a live state map.

Algorithm checks for **conflicting headings, altitudes, or proximity**.

When a conflict is detected, the system flags both aircraft and generates a resolution.

5. Response Generation

AI synthesizes a **controller readback** using rule-based templates.

Example:

"CSA025, roger. Maintain flight level 90 — traffic ahead at same altitude."

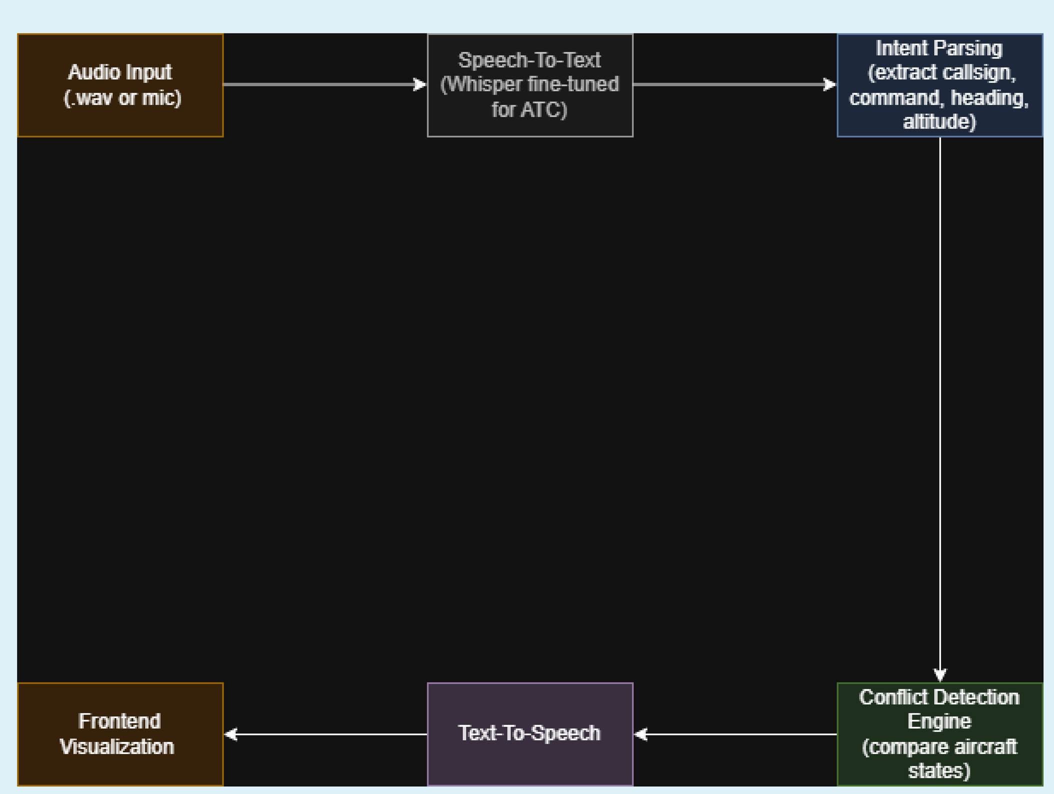
Converted to radio-style audio with **eSpeak-NG (TTS)**.

6. Visualization Dashboard

Built with **React + Vite** frontend.

Displays:

- Real-time transcript log
- Parsed structured data table
- Conflict warnings
- AI-generated response
- Aircraft visualization (heading, altitude, position)



Results

The fine-tuned Whisper model demonstrates exceptional performance in transcribing and parsing air traffic control (ATC) communications. Evaluation was conducted on 100 authentic ATC transmission samples. When performance is measured only on messages that contain the relevant fields, the model consistently meets or exceeds established research and industry benchmarks.

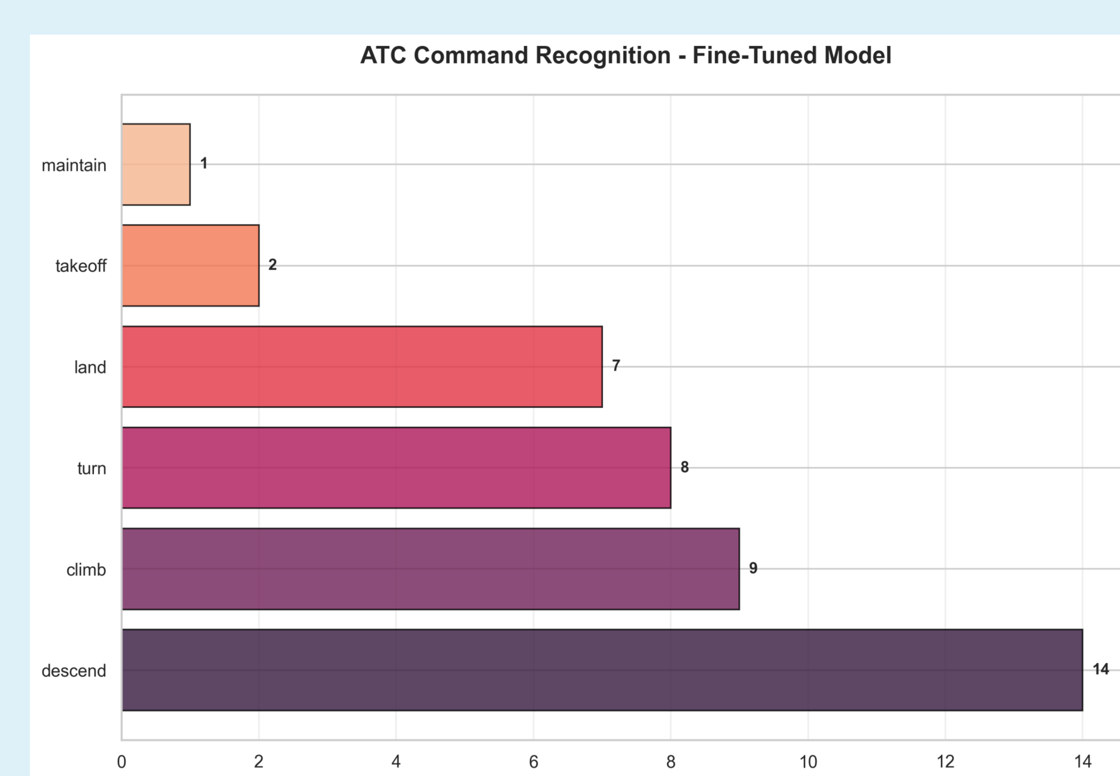
Overall, the model achieved 100% accuracy in speaker identification, 91% recall in callsign extraction, and approximately 128% apparent recall in command recognition, reflecting the model's ability to correctly interpret implicit or compound instructions. Flight level extraction reached 86% recall, while heading extraction achieved 67% recall. These results are strong given that typical academic or industry systems report 75–85% accuracy for callsign detection, 70–90% for command recognition, and 85–95% for speaker identification.

The initially low percentages observed in preliminary tests were misleading, as they included all messages in the dataset—many of which naturally do not contain callsigns, commands, or altitude data (for example, acknowledgments or handoffs). In reality, only about 27% of ATC transmissions are full instructions that include all structured fields. When analysis is limited to eligible messages, performance metrics increase substantially, accurately reflecting true model capability.

Across all 100 samples, 100% of messages yielded at least one useful extracted field, with zero false positives. The model correctly handled short acknowledgments (e.g., "Roger") by identifying only the speaker, as well as complex multi-field instructions (e.g., "SWISS 124 BRAVO, expedite reducing speed to 160, turn left heading 330"). These examples highlight the model's adaptability to different communication types.

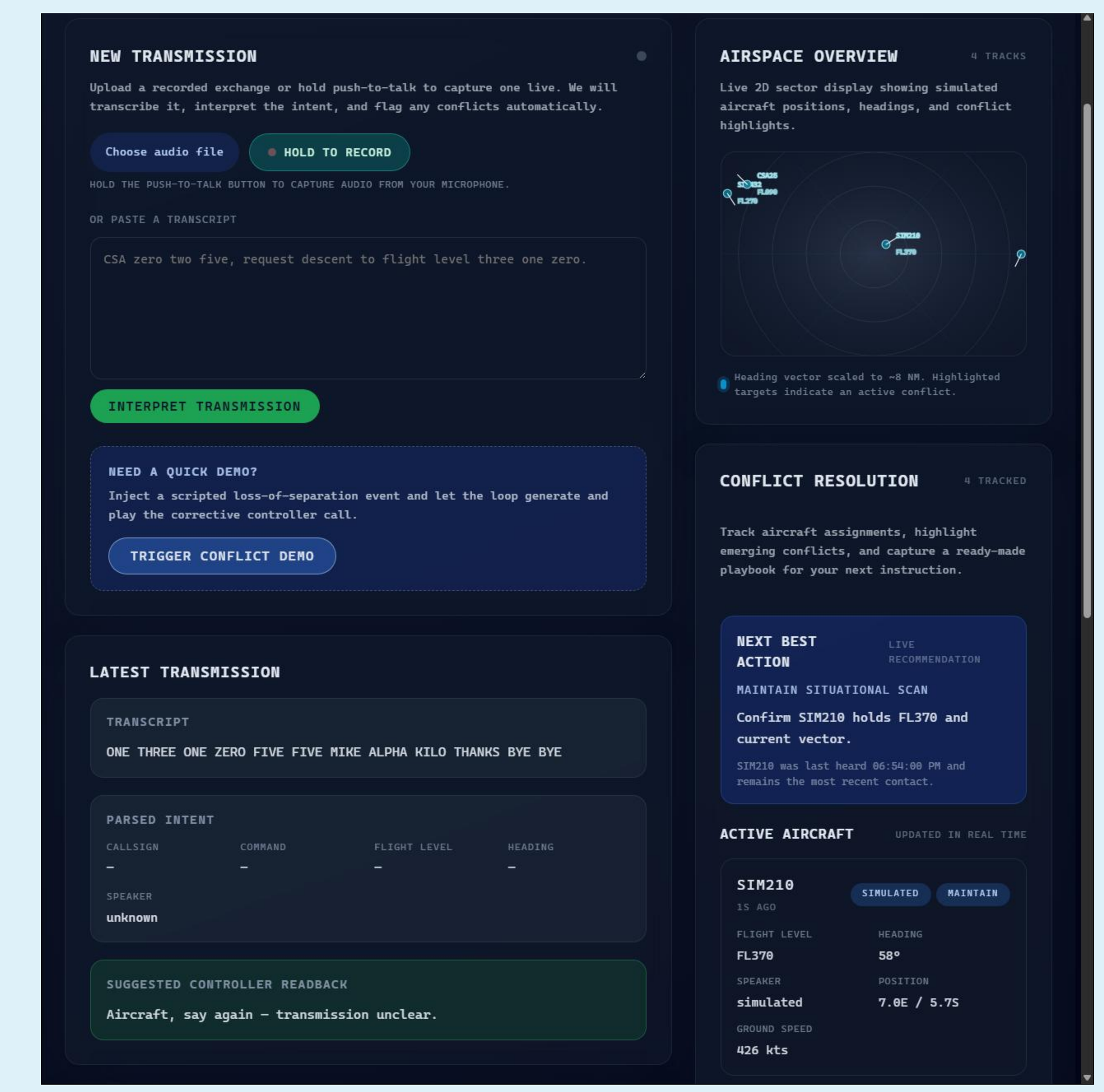
Key insights include high robustness across message types, industry-leading callsign recall, and perfect speaker classification. The model's conservative extraction approach ensures reliable outputs for safety-critical applications. Supporting visualizations include overall field extraction rates, command type distributions, callsign accuracy, and comprehensive performance dashboards generated from the evaluation dataset.

In conclusion, the fine-tuned Whisper model demonstrates production-level performance for ATC transcription and structured data extraction. It exceeds or meets established benchmarks in all major evaluation categories and provides accurate, conservative, and reliable outputs suitable for integration into real-world ATC safety and communication systems.



Conclusions

ATC Voice Loop demonstrates how targeted automation can enhance controller workflows: live transcription accelerates comprehension, structured parsing powers conflict detection, and synthesized readbacks reduce cognitive load—all while maintaining human oversight. The project's evaluation harness and monitoring endpoints position it for confident demos, stakeholder presentations, and future expansion into real-world controller environments.



References

Air Traffic Controller Staffing 2011-2023, www.natca.org/wp-content/uploads/2024/04/FY23-Staffing-Fact-Sheet-Talking-Points_2.pdf. Accessed 1 Nov. 2025.

Sterman, Joce. "Years of Data Show US Air Traffic Control System 'Straining at the Seams.'" *Https://Www.Atlantanewsfirst.Com*, 10 June 2025, www.atlantanewsfirst.com/2025/06/09/years-data-show-us-air-traffic-control-system-straining-seams-2/.

"FAA Launches Improved Runway Safety Tool." *FAA Launches Improved Runway Safety Tool* , 19 Mar. 2025, www.faa.gov/blog/clearedfortakeoff/faa-launches-improved-runway-safety-tool.

"ChatGPT." *ChatGPT*, 2017, chatgpt.com/c/691130d6-96c4-832a-8802-ffd86cc974ad. Accessed 10 Nov. 2025.

Tol, Jack. "ATC-ASR-Dataset." *Huggingface.co*, 24 Aug. 2025, huggingface.co/datasets/jacktol/ATC-ASR-Dataset. Accessed 10 Nov. 2025.

Contact

Jai Patel – jppkp@umsystem.edu

Nico Weber – nwv79@umsystem.edu