BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2026
http://excel.fit.vutbr.cz

Automatic filling of electronic flight strips

Bc. Veronika Nevařilová*

Abstract
Air traffic controllers at small aerodromes maintain a book of flights — a record of all flights that passed through their airspace, where each flight can be represented by a so called *strip*, containing all important information about it. Traditionally kept on paper, many aerodromes are now transitioning to digital book of flights. But even when using **digital flight strips**, controllers must fill them manually based on the communications with pilots. This paper presents an **automated pipeline** tailored for a future digital book of flights at Kunovice airport (LKKU) that **processes air traffic communication** recordings and extracts the required information with little to **no human intervention** needed.
The pipeline consists of five sequential stages: **preprocessing, automatic speech recognition, data extraction, speaker identification and postprocessing**, and is able to process audio segments on the order of seconds. When tested against a full day of air traffic recordings from LKKU with no human correction, the pipeline achieved an end-to-end F1 score of 72.8%, where a record is counted as correct only if all fields exactly match the ground truth. With manual correction of callsign assignments, F1 rose to 80.2%.
By automating flight book data entry, the pipeline could reduce the administrative burden on controllers at small aerodromes. Also, the approach generalises to any airport and could be adapted to other aerodromes or environments where manual logging remains standard practice.

*xnevar00@stud.fit.vutbr.cz, *Faculty of Information Technology, Brno University of Technology*

1. Introduction

Smaller aerodromes like Kunovice (LKKU) are required to maintain a book of flights — a chronological record of flights through their airspace. The flights can be represented in the book by so called *strips* (Fig. 1) that contain all important information about them. Besides serving as a history record, they also help air traffic controllers track active traffic and maintain their situational awareness. At Kunovice and many other airports this record is still kept on paper, meaning any downstream use of the data (invoicing landing fees or compiling statistics) requires the controller or administrative staff to manually retype it into a computer. Digitalizing the book of flights and automating the process of the strip filling would reduce the risk of omissions, lower the workload of the controller or any other person processing it, and produce a digital record directly usable for further processing.

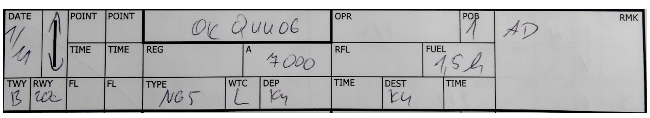

Figure 1. Example of paper flight strip at Kunovice.

The core goal was to create a system that, given a set of chronological audio recordings of air traffic radio communication, automatically extracts structured data ready to be processed by a digital book of flights. This extracted data would include callsign, aircraft type, departure and destination airport, altitudes and runway assignment among other things.

Since the only input of the system is audio, the core components needed are automatic speech recognition, speaker identification, and structured information extraction. The aviation field presents unique challenges for all three: transmissions are noisy and follow domain-specific phraseology that general-purpose systems handle poorly. Each component therefore requires some degree of domain-specific adaptation.

2. Proposed system

The input audio is processed sequentially through the following five stages:

1. **Preprocessing block** – segmenting the incoming audio into individual transmissions based on silence detection, which proved more reliable

than general-purpose diarization models for this type of input.

2. **ASR block** – transcribing speech into text using NeMo Parakeet model [1] finetuned on $\sim$4 hours of czech and english LKKU air traffic audio recordings [2].
3. **Information extraction block** – using an LLM [3] to extract structured data from the transcribed speech, combining semantically-driven extraction with rule-based normalization.
4. **Speaker identification block** – determining whether the current transmission belongs to a known speaker from the set of active aircraft or to a new one, using a ResNet-152 speaker recognition model adapted for low-bandwidth speech [4]. The decision strategy is based on a combination of voice embedding similarity and callsigns mentioned in the transmission.
5. **Output block** – gathering data from the individual blocks and creating the final JSON output ready to be processed by the book of flights.

3. Results

The pipeline was evaluated on LKKU data both at the level of individual blocks and as a complete end-to-end system.

3.1 Individual blocks

The preprocessing block achieved 97.5% accuracy in segmenting the audio stream into individual transmissions. The ASR block reached word accuracy of 87.2%. The information extraction block achieved 98.8% field-level F1 score in extracting structured information from transcribed text. Speaker identification block was evaluated on a dedicated test set and achieved 100% accuracy.

3.2 End-to-end evaluation

The complete pipeline was evaluated on one full day of air traffic recordings, comprising 396 transmissions totalling 30 minutes of audio. Performance was measured using end-to-end F1 score, where a prediction is counted as correct only if all fields exactly match the ground truth. Without any human intervention, the pipeline achieved 72.8% F1. When incorrectly extracted callsigns were manually corrected (4 corrections out of almost 400 segments), F1 rose to 80.2%.

The impact of 4 corrections on the accuracy reflects the cascading nature of errors in the pipeline: a single mistake early in the session causes all subsequent outputs for that speaker to be attributed incorrectly. A single correction therefore has an outsized positive effect, while a single uncorrected error can affect a large portion of the results. However, the option to correct the system is consistent with its intended use — the outputs are visible in the book of flights in real time and can be corrected as they appear.

Notably, two bottlenecks emerged as the main sources of error. The first is the combination of ASR and LLM extraction: when the ASR misrecognises a word, the LLM receives a corrupted transcript and may extract incorrect or missing values as a result. The second is speaker identification and automatic callsign assignment, which remain the most challenging aspects of the system.

4. Conclusions

The pipeline demonstrated reliable performance across all stages, achieving **80% accuracy with minimal human correction**. Beyond digital flight strip filling, the pipeline has potential for a much broader range of applications. By capturing not only what was said, but who said it and which aircraft it concerns, it can serve as a foundation for virtually any downstream processing of aviation communication, e.g. traffic analysis or training and simulation, by replaying real communications for controller or pilot training.

Future work could improve performance further by collecting and annotating more data to fine-tune the ASR model, which would directly reduce errors propagating into the information extraction stage. Additionally, processing audio with awareness of previous transmissions in the session would allow the system to resolve ambiguities using conversational context. Both bottlenecks — ASR-LLM error propagation and speaker identification — would benefit from this approach.

Acknowledgements

I would like to thank my supervisor Ing. Igor Szőke, Ph.D for his invaluable help and guidance throughout the whole development.

References

[1] Monica Sekoyan, Nithin Rao Koluguri, Nune Tadevosyan, et al. Canary-1B-v2 & Parakeet-TDT-0.6B-v3: Efficient and High-Performance Models for Multilingual ASR and AST, 2025.

[2] Veronika Nevařilová. Automatický přepis řeči letecké komunikace do textu. Bakalářská práce, Vysoké učení technické v Brně, Fakulta informačních technologií, Brno, 2024. Vedoucí práce Ing. Igor Szőke, Ph.D.

[3] Gheorghe Comanici, Eric Bieber, Mike Schaekermann, Ice Pasupat, Noveen Sachdeva, et al. Gemini

2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities, 2025.

[4] Sara Barahona, Anna Silnova, Ladislav Mošner, Junyi Peng, Oldřich Plchot, et al. Analysis of ABC Frontend Audio Systems for the NIST-SRE24. In *Interspeech 2025*, pages 5763–5767, 2025.