Scribe NLP: Unleashing the potential of ATC Voice Communication

Fabien Betouret Candau,[1] Jeremi Carol,[1] Sylvain Martin[1] and Thomas Roques[1]

[1] *Direction de la Technique et de l'Innovation, Direction des Services de la Navigation Aérienne,*
Direction Générale de l'Aviation Civile, Toulouse 31100, France
[†]*email: scribe.dsna-dti@aviation-civile.gouv.fr*

Scribe is an automated natural language processing (NLP) service tailored to the French air traffic control (ATC) environment. Used by numerous DSNA innovative projects, Scribe enhances each step of an air traffic controllers work: upstream with the automation of simulations during their training, in real-time by adding a voice-based safety net and integrated assistance tools within the controller HMI, and post-operation during safety analysis transcription tasks.

Key Words: ATC, NLP, ASR, ATM, STT, TTS, NLU, NER

1. Introduction

Air traffic controllers play a crucial role in ensuring the safety and efficiency of air traffic. While radar and flight plan data are widely used in various tools and services, voice communications' potential remains untapped due to their analog nature. Since 2000, research projects have aimed to extract valuable information from voice communications but have not reached widespread usage due to computational limitations and lack of generalization.

ATC speech processing faces challenges like noisy channels, fast speech rates, diverse accents, and multilingual environments including code switching. Additionally, formal phraseology is not always strictly followed, especially in native language usage. This complex nature of ATC environment makes off-the-shelf Automatic Speech Recognition (ASR) engines inadequate for reliable services.

2. Scribe ATC Natural Language Processing

Fig. 1. Scribe logo.

In 2021, DSNA started the Scribe initiative to build a specialized natural language processing (NLP) system for ATC, inspired by the recent success of multiple deep learning architectures in this field of research. The aim was to deploy a generic service that could be used in various tools, services, and safety nets designed to benefit air traffic controllers and the broader ATC staff.
This service, able to process both live frequency streams and offline recordings, provides the following features:
- Transcription: a multilingual Speech-To-Text (STT) service that covers all flight phases (Tower, Approach, and En-Route). The latest version achieves a Word Error Rate (WER) of 7.79% and a Character Error Rate (CER) of 5.22%, with a test set of unseen voices from operational exchanges recorded in 11 DSNA centers.
- Understanding: a Named Entity Recognition (NER) service that converts plain text into structured data (based on SESAR 2020 ontology[1]). It extracts the callsign, the intent (clearance, readback, etc.) and the message content (command, value, unit, qualifiers) of En-Route and Approach phases for the time being.
- Callsign Detection: a real-time algorithmic service for detecting callsigns in live streams.
- Voice Synthesis: multi-lingual multi-accent Text-To-Speech (TTS) service customizable to accommodate ATC specificities.

3. Applications

DSNA has already developed five innovative projects on top of Scribe Services:
- *ELVIRA*, DSNA's safety analysis tool, embeds Scribe transcription service since November 2022 to assist safety teams during the tedious task of transcription of safety related events vocal exchanges.
- *CROSSCHECK* is a safety net that uses voice recognition to issue real-time alerts when it detects incorrect pilot read-backs or erroneous controller clearances.
- *PILOTE-VIRTUEL* enhances the pseudo-pilot capacity during ATC simulations by detecting in real-time the aircraft referred to by a control clearance, identifying the instructions given (type, value, unit, etc.), and automatically reading them back using voice synthesis.
- *SPOT* highlights the calling aircraft in real-time on the controller's HMI upon initial contact, then animates flight label on each subsequent readback or request.
- *SELECT* automatically inputs controller clearances into flight labels within stripless ATC HMI.

References

1) 1) H. Helmke et al., "Ontology for Transcription of ATC Speech Commands of SESAR 2020 Solution PJ.16-04", 2018, pp. 1-10, doi: 10.1109/DASC.2018.8569238.