

---

# **System Requirements Specification**

**for**

## **RTube Kaldi Team**

**Version 1.0 approved**

**Prepared by Tabitha O'Malley**

**CS 490 Kaldi Team Fall 2023**

**09/29/2023**

Table of Contents

Table of Contents	Error! Bookmark not defined.
Revision History	3
1. Introduction	Error! Bookmark not defined.-2
1.1	Purpose
1.2Document	Conventions
1.3Intended Audience and	Error! Bookmark not defined.
1.4Product	Reading Suggestions
1.5	Error! Bookmark not defined.
2. Overall Description	Scope
2.1Product	References
2.2Product	Error! Bookmark not defined.
2.3User Classes and	Error! Bookmark not defined.
2.4Operating	Perspective
2.5Design and Implementation	Error! Bookmark not defined.
2.6User	Functions
2.7Assumptions and	Error! Bookmark not defined.
3. External Interface Requirements	Characteristics
3.1User	Error! Bookmark not defined.
3.2Hardware	Environment
3.3Software	Error! Bookmark not defined.
3.4Communications	Constraints
4. System Features	Error! Bookmark not defined.
4.1System Feature	Documentation
4.2System Feature 2 (and so on)	Error! Bookmark not defined.
5. Other Nonfunctional Requirements	Dependencies
	Error! Bookmark not defined.
	Error! Bookmark not defined.
	Error! Bookmark not defined.
	Interfaces
	4
	Interfaces
	Error! Bookmark not defined.
	Interfaces
	Error! Bookmark not defined.
	Interfaces
	Error! Bookmark not defined.
	Error! Bookmark not defined.
	Error! Bookmark not defined.
	1
	Error! Bookmark not defined.
	Error! Bookmark not defined.
	Error! Bookmark not defined.

5.1Performance	Requirements	7
5.2Safety	Requirements	Error! Bookmark not defined.
5.3Security	Requirements	Error! Bookmark not defined.
5.4Software	Quality	Attributes
5.5Business	Rules	Error! Bookmark not defined.
6. Other Requirements		Error! Bookmark not defined.
Appendix A: Glossary		Error! Bookmark not defined.
Appendix B: Analysis Models		Error! Bookmark not defined.
Appendix C: To Be Determined List		Error! Bookmark not defined.

Revision History

Name	Date	Reason For Changes	Version
All	09/29/23	Start the document	V1

# 1. Introduction

## 1.1 Purpose

*<Identify the product whose software requirements are specified in this document, including the revision or release number. Describe the scope of the product that is covered by this SRS, particularly if this SRS describes only part of the system or a single subsystem.>*

The RTube web application provides an interface for users to track the flight paths of aircrafts, and to listen to live ATC tower radio chatter with the option to transcribe speech to text in real time. The live speech-to-text transcription is performed using an ASR model developed using the Kaldi ASR toolkit.

The purpose of the model is to transcribe live spoken communications between aircraft pilots and ATC operators into text that shall be displayed in real-time by the RTube web application. The model in question was created by modifying previously existing Kaldi-developed ASR models with a 30-hour training dataset tailored for aviation communication. The dataset consists of radio transmissions between aircraft pilots and ATC from the United States and Europe.

## 1.2 Document Conventions

**Times New Roman font, 11 point** is the primary body of the document under each section/subsection header.

**Times New Roman font, 18 point, bold** is used for section headers.

**Times New Roman font, 14 point, bold** is used for subsection headers.

**Yellow highlights** are indicative of pending review by another team member.

**Italicized Text placed within <>** are indicative of what the section is supposed to detail.

**Bold Text within the document body** is used to indicate any important terms or ideas that need to stand out from the standard text.

**Terms or Ideas at the same Indent level** indicate the same level of importance.

**Terms or Ideas at the different Indent level** indicate the differing levels of importance, or a hierarchy.

## 1.3 Intended Audience and Reading Suggestions

*<Describe the different types of reader that the document is intended for, such as developers, project managers, marketing staff, users, testers, and documentation writers. Describe what the rest of this SRS contains and how it is organized. Suggest a sequence for reading the document, beginning with the overview sections and proceeding through the sections that are most pertinent to each reader type.>*

The intended audience for the RTube web application is aircraft pilots and ATC operators. One of the more specific applications for RTube shall be for ERAU student pilots and instructors to evaluate communications between the student and ATC. This document contains a detailed description of the product and the requirements given by the customer. The first section of this document shall introduce the project and give a basic understanding of its purpose. The next section shall give an overall description of the project. Section 3 shall provide the technical aspect of the project, typically used during testing and replication of the product. Sections 4 and 5 shall provide the requirements and features of the project.

Commented [1]: Milan

Commented [2]: This needs to be rewords to fit just our section of the project

Commented [3]: aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

Commented [4]: aaaaaaaa?

To improve understanding of this document and the project itself, reading documentation related to ASR, NLP, and Kaldi is highly recommended. Further reading includes documentation regarding phonetics and aviation phraseology.

## 1.4 Product Scope

*<Provide a short description of the software being specified and its purpose, including relevant benefits, objectives, and goals. Relate the software to corporate goals or business strategies. If a separate vision and scope document is available, refer to it rather than duplicating its contents here.>*

As an aircraft pilot, learning to communicate with ATC is a daunting task. Having to maneuver an aircraft while also relaying important information back and forth to ATC in a clear and concise manner is a skill not for the faint of heart. A few methods to practice this skill exist, one of which is listening to LiveATC, a website that provides live audio streams of various airports and flights from around the world. However, this can prove difficult without the sufficient knowledge of aviation phraseology. Despite it being designed to mitigate miscommunication, the phraseology is highly intricate and requires at least hundreds of hours of practice to learn its idiosyncrasies. Due to this, many new aircraft pilots are required to attend special training simply to learn how to properly communicate with ATC - a testament to the task's complexity.

The goal of this project is to transcribe the complex communications between ATC and aircraft pilots into text and have it displayed live via the RTube web application . By transcribing speech to text in real time, the student pilots can dramatically reduce the special training required to simply understand spoken aviation phraseology. In essence, students can be able to learn outside of a classroom and without an instructor present. Secondly, by having a live map and a real-time transcription, students can better understand the different contexts that specific phraseology is used in.

## 1.5 References

*<List any other documents or Web addresses to which this SRS refers. These may include user interface style guides, contracts, standards, system requirements specifications, use case documents, or a vision and scope document. Provide enough information so that the reader could access a copy of each reference, including title, author, version number, date, and source or location.>*

Kaldi Team Product Vision Statement. Kaldi Team. 19 September 2023. Kaldi Drive. [Kaldi Product Vision Statement](#)

<https://www.kaldi-asr.org/doc/about.html>  
<https://www.liveatc.net/faq/>  
<https://pandas.pydata.org/about/>  
<https://openpyxl.readthedocs.io/en/stable/>

Commented [5]: should this be same as SDS?

Commented [6]: it it the reference for this document so it may be similar but it not the same

## 2. Overall Description

### 2.1 Product Perspective

Commented [7]: Tabitha

<Describe the context and origin of the product being specified in this SRS. For example, state whether this product is a follow-on member of a product family, a replacement for certain existing systems, or a new, self-contained product. If the SRS defines a component of a larger system, relate the requirements of the larger system to the functionality of this software and identify interfaces between the two. A simple diagram that shows the major components of the overall system, subsystem interconnections, and external interfaces can be helpful.>

Utilizing results from other ASR teams that have used Kaldi for ATC, RTube shall fill the need for ASR-based training for aeronautical phraseology. This project shall focus on the modeling for RTube by learning and adjusting the models to fit the application.

ADD System Configuration Diagram V2

### 2.2 Product Functions

Commented [8]: Milan

<Summarize the major functions the product must perform or must let the user perform. Details shall be provided in Section 3, so only a high level summary (such as a bullet list) is needed here. Organize the functions to make them understandable to any reader of the SRS. A picture of the major groups of related requirements and how they relate, such as a top level data flow diagram or object class diagram, is often effective.>

The ASR model shall accurately transcribe ATC communications spoken in General American English into text that is displayed via the RTube web application.

### 2.3 User Classes and Characteristics

Commented [9]: Adam Gallub

<Identify the various user classes that you anticipate shall use this product. User classes may be differentiated based on frequency of use, subset of product functions used, technical expertise, security or privilege levels, educational level, or experience. Describe the pertinent characteristics of each user class. Certain requirements may pertain only to certain user classes. Distinguish the most important user classes for this product from those who are less important to satisfy.>

The quick brown fox jumped over the lazy dog (placeholder).

## 2.4 Operating Environment

Commented [10]: Adam Gallub

*<Describe the environment in which the software shall operate, including the hardware platform, operating system and versions, and any other software components or applications with which it must peacefully coexist.>*

The ASR model shall operate in the RTube web application using the Flask web framework. Flask allows the software to interact with the Python libraries *pandas* and *openpyxl*; the former is used for data analysis and manipulation, while the latter is used for reading file types used by Microsoft Excel (e.g., .xlsx, .xlsm, .xltx, .xltm).

The ASR model shall obtain information from radio communication systems, flight radar applications, and ATC channels, similar to the likes of FlightRadar24. The difference is that RTube shall be able to transcribe speech to text.

## 2.5 Design and Implementation Constraints

A significant portion of the data used to create the ASR model utilized by RTube came from outside of the United States - mainly Europe. The data consists of both the English language and the local vernacular of each respective country. Despite the different varieties of spoken English accessible to the ASR model, the model is constrained to General American English. This is due to RTube itself being geographically restricted to the state of Florida. In addition, RTube only has access to the Daytona Beach Airport ATC radio chatter. Other languages, as well as other varieties of the English language (e.g., British English, General Australian English, Indian English, etc.) are not guaranteed to be recognized by the ASR model.

## 2.6 User Documentation

The ASR model has no user documentation, as the user shall never directly interact with it. Instead, the user interacts with the RTube web application.

## 2.7 Assumptions and Dependencies

The ASR model shall be created under a multitude of assumptions and dependencies. The assumptions are as follows: there is clear speech from both aircraft pilots and ATC; there is low background interference (e.g., static); all ATC radio transmissions are exclusively obtained from Daytona Beach Airport; the RTube web application is constrained to the state of Florida; and that 30 hours of training data is sufficient to produce accurate transcriptions. The dependencies are as follows: all audio input is General American English; and the NeMo team creates a function web application and user interface (RTube).

# 3. External Interface Requirements

## 3.1 User Interfaces

This project does not have any user interfaces.

### 3.2 Hardware Interfaces

*<Describe the logical and physical characteristics of each interface between the software product and the hardware components of the system. This may include the supported device types, the nature of the data and control interactions between the software and the hardware, and communication protocols to be used.>*

This project does not have any hardware interface requirements.

### 3.3 Software Interfaces

The ASR model created in Kaldi interacts with the NeMo project, which acts as an interface for the user to interact with the Kaldi software. There shall be connections with ATC networks to read audio signals and send text to be displayed on the NeMo side (RTube). To train the AI, a 30-hour ATC communications database shall be used. There shall be a database of phones and triphones for training and transcription.

### 3.4 Communications Interfaces

*<Describe the requirements associated with any communications functions required by this product, including e-mail, web browser, network server communications protocols, electronic forms, and so on. Define any pertinent message formatting. Identify any communication standards that shall be used, such as FTP or HTTP. Specify any communication security or encryption issues, data transfer rates, and synchronization mechanisms.>*

The ASR model created in Kaldi interacts with the NeMo Project in the form of a web application called RTube. The communication functions shall also work within a frequency channel of all Florida ATC frequencies. The Kaldi software also pulls from previous data collected. Since RTube is a web application, the communications standard in place is HTTP.

Commented [11]: Adam Gallub



## 4. System Features -

Commented [12]: Adam Gallub

*<This template illustrates organizing the functional requirements for the product by system features, the major services provided by the product. You may prefer to organize this section by use case, mode of operation, user class, object class, functional hierarchy, or combinations of these, whatever makes the most logical sense for your product.>*

*-None at the moment, however there shall be, didn't know if I should've put future system features or not*

### 4.1 System Feature 1

*<Don't really say "System Feature 1." State the feature name in just a few words.>*

#### 4.1.1 Description and Priority

*<Provide a short description of the feature and indicate whether it is of High, Medium, or Low priority. You could also include specific priority component ratings, such as benefit, penalty, cost, and risk (each rated on a relative scale from a low of 1 to a high of 9).>*

#### 4.1.2 Stimulus/Response Sequences

*<List the sequences of user actions and system responses that stimulate the behavior defined for this feature. These shall correspond to the dialog elements associated with use cases.>*

#### 4.1.3 Functional Requirements

*<Itemize the detailed functional requirements associated with this feature. These are the software capabilities that must be present in order for the user to carry out the services provided by the feature, or to execute the use case. Include how the product should respond to anticipated error conditions or invalid inputs. Requirements should be concise, complete, unambiguous, verifiable, and necessary. Use "TBD" as a placeholder to indicate when necessary information is not yet available.>*

*<Each requirement should be uniquely identified with a sequence number or a meaningful tag of some kind.>*

REQ-1:

REQ-2:

### 4.2 System Feature 2 (and so on)

## 5. Other Nonfunctional Requirements

### 5.1 Performance Requirements

*<If there are performance requirements for the product under various circumstances, state them here and explain their rationale, to help the developers understand the intent and make suitable design choices. Specify the timing relationships for real time systems. Make such requirements as specific as possible. You may need to state performance requirements for individual functional requirements or features.>*

Commented [13]: tabitha

### 5.2 Safety Requirements

*<Specify those requirements that are concerned with possible loss, damage, or harm that could result from the use of the product. Define any safeguards or actions that must be taken, as well as actions that must be prevented. Refer to any external policies or regulations that state safety issues that affect the product's design or use. Define any safety certifications that must be satisfied.>*

Commented [14]: Tabitha

Filter out words that are not in the lexicon (e.g., someone says Emory-Riddle instead of Embryo-Riddle)

Lexicon contains words that are standardized

### 5.3 Security Requirements

*<Specify any requirements regarding security or privacy issues surrounding use of the product or protection of the data used or created by the product. Define any user identity authentication requirements. Refer to any external policies or regulations containing security issues that affect the product. Define any security or privacy certifications that must be satisfied.>*

Commented [15]: Adam Gallub

In order to utilize the ASR model within RTube, users shall require a registered account to access the web application. Proper login credentials, in the form of a unique username and a password, shall be prompted for users to input into RTube before being granted access. No further authentication is required to access RTube (e.g., MFA). If a user does not have a registered RTube account, they shall be able to create a new account (and the required credentials) before accessing the web application. Users without registered accounts cannot access RTube, and thus the ASR model utilized by the web application.

## 5.4 Software Quality Attributes

Commented [16]: Adam Gallub

*<Specify any additional quality characteristics for the product that shall be important to either the customers or the developers. Some to consider are: adaptability, availability, correctness, flexibility, interoperability, maintainability, portability, reliability, reusability, robustness, testability, and usability. Write these to be specific, quantitative, and verifiable when possible. At the least, clarify the relative preferences for various attributes, such as ease of use over ease of learning.>*

The software shall retrieve the location of an aircraft in the air in real time, and display each aircraft's flight pattern. The software shall update and save text from the time the aircraft takes off to when it lands, so that the user can look at the text channels.

## 5.5 Business Rules

Commented [17]: Adam Gallub

*<List any operating principles about the product, such as which individuals or roles can perform which functions under specific circumstances. These are not functional requirements in themselves, but they may imply certain functional requirements to enforce the rules.>*

RTube users shall have any aircraft in the sky within the state of Florida displayed on-screen, and have the following information displayed: model, callsign, flight path, destination, and live communication between ATC and aircraft pilots transcribed in real time.

## 6. Other Requirements

<Define any other requirements not covered elsewhere in the SRS. This might include database requirements, internationalization requirements, legal requirements, reuse objectives for the project, and so on. Add any new sections that are pertinent to the project.>

### Appendix A: Glossary

<Define all the terms necessary to properly interpret the SRS, including acronyms and abbreviations. You may wish to build a separate glossary that spans multiple projects or the entire organization, and just include terms specific to a single project in each SRS.>

ASR: Automatic speech recognition  
ATC : Air Traffic Control  
ERAU: Embry-Riddle Aeronautical University  
MFA: Multi-factor authentication  
NLP: Natural language processing

Kaldi: a toolkit for speech recognition written in C++ and licensed under the Apache License v2.0

Phone: a basic unit of spoken language  
Phoneme: a combination of phones  
Triphone: a combination of phonemes

British English: an umbrella term for the varieties of English spoken in Great Britain (England, Scotland, Wales)  
General American English: an umbrella term for the variety of American English spoken by the majority of Americans; i.e., the stereotypical American accent.  
General Australian: an umbrella term for the varieties of English spoken in Australia  
Indian English: an umbrella term for the varieties of English spoken in India

### Appendix B: Analysis Models

<Optionally, include any pertinent analysis models, such as data flow diagrams, class diagrams, state-transition diagrams, or entity-relationship diagrams.>

The quick brown fox jumped over the lazy dog (placeholder).

### Appendix C: To Be Determined List

<Collect a numbered list of the TBD (to be determined) references that remain in the SRS so they can be tracked to closure.>

The quick brown fox jumped over the lazy dog (placeholder).

Commented [18]: Everybody

Commented [19]: should this be same as SDS or no? many terms are probably going to be same

Commented [20]: like the references it will be similar but not the same