**Case Study - Requirements Specification Document**

**1. Abstract:**

This document presents the software requirements specification (SRS) for an automatic subtitle generator for DaVinci Resolve. The system makes use of OpenAI Whisper for speech-to-text transcription and has an integrated user interface within DaVinci Resolve for generating, editing, and applying subtitles to video timelines. The document is presented according to the IEEE standard for a requirements specification document.

**2. Introduction:**

**2.1 Purpose:**

The intent of this document is to outline the external requirements for an auto subtitle generation system to operate within DaVinci Resolve. It outlines what inputs are expected, what outputs to provide, constraints on the system, and functional requirements.

**2.2 Scope:**

This document formalizes the system requirements and is used as a reference by the developers to facilitate proper implementation. Any subsequent change to the requirements should undergo a formal change approval process. The developer should clarify and should not make changes in specifications without approval from the client.

**2.3 Definitions, Acronyms, Abbreviations:**

- Whisper: OpenAI’s speech-to-text model.

- DaVinci Resolve: A professional video editing software by Blackmagic Design.

- UI: User Interface.

- Transcription: The process of converting speech into text.

- Subtitle File: A text file containing time-synchronized captions.

**2.4 References:**

Not applicable.

**2.5 Developer’s Responsibilities:**

The developer has the following responsibilities:  
(a) Deployment of the auto subtitle generator.  
(b) DaVinci Resolve integration with the system.  
(c) Installing the software and its compatibility with various operating systems.  
(d) Carrying out user training whenever necessary.  
(e) System maintenance for one year after installation.

**3. General Description:**

**3.1 Product Functions Overview:**

The system enables subtitles to be generated from dialogue in video files by:

- Inspecting the portion of the DaVinci Resolve video timeline chosen.

- Utilizing OpenAI Whisper to translate speech into text.

- Discovering all supported languages for OpenAI Whisper, translating them to English, and producing accurate English subtitles.

- Producing English subtitles from audio that is not English.

- Executing dialect separation to separate speakers.

- Creating a subtitle file with appropriate timecodes.

- Providing subtitle placement and style adjustment for the users.

- Exporting the subtitles in several formats (e.g., SRT, VTT, DaVinci Resolve Text+ format).

**3.2 User Characteristics:**

The primary users are video editors and content creators using DaVinci Resolve. Users are expected to have basic proficiency in video editing but may not be technically inclined toward coding or automation tools.

**3.3 General Constraints:**

The system must:

- Run on Windows, macOS, and Linux.

- Be compatible with DaVinci Resolve’s scripting API.

- Process subtitle generation within a reasonable time (<1 minute for a 5-minute clip).

**3.4 General Assumptions and Dependencies:**

- The user has DaVinci Resolve installed.

- The system requires an active internet connection for Whisper API usage.

**4. Specific Requirements:**

**4.1 Inputs and Outputs:**

Inputs:

1. Chosen video clip from DaVinci Resolve.

2. Customizable parameters (e.g., subtitle position, subtitle font and size, output type).

Outputs:

1. Subtitle file (SRT, VTT, or DaVinci Resolve Text+ format).

2. Preview transcription in the UI.

3. Error messages for incorrect inputs or processing failures.

**4.2 Functional Requirements:**

1. Create subtitles by transcribing dialogue from a chosen video clip.

- Inputs: Video clip, user settings.

- Outputs: Subtitle file.

2. Identify and translate all supported languages to English subtitles.

- Inputs: Audio of the video.

- Outputs: English subtitle file.

3. Conduct dialect separation to distinguish speakers.

- Inputs: Audio input with multiple speakers.

- Inputs: Transcription generation.

- Outputs: Speaker-separated subtitle file.

4. Show transcription preview in the UI.

- Inputs: Transcription generated.

- Outputs: Editable text interface.

5. Provide users with options to modify subtitle settings prior to export.

- Inputs: User-controlled settings.

- Outputs: Modified subtitle file.

6. Manage errors and print relevant messages.

- Inputs: System errors, wrong inputs.

- Outputs: Error logs and alerts.

**4.3 External Interface Requirements:**

User Interface:

- A graphical interface integrated into DaVinci Resolve with buttons for processing subtitles.

- File selection options for loading/exporting subtitle files.

**4.4 Performance Constraints:**

- Subtitle creation for a 5-minute video should be finished within 1 minute.  
- The system must be run without introducing too much lag into DaVinci Resolve.

**4.5 Design Constraints:**

Software Constraints:

- The system should be compliant with DaVinci Resolve's Python API.

- The system should employ OpenAI Whisper as the transcription utility.

Hardware Constraints:

- The system should execute on machines of at least 8GB RAM and multi-core processor.

Acceptance Criteria:

- The system should be able to generate subtitles for at least 90% of test scenarios with decent accuracy.

- DaVinci Resolve integration should be seamless, no crashes or substantial UI issues.