

Revised System Requirements Specification

1. Introduction

1.1 Purpose

This document outlines the updated requirements for the Textify system. Textify aims to leverage state-of-the-art AI technologies for text recognition, translation, and related features. The primary features of Textify include recognizing text from images using Optical Character Recognition (OCR), translating text between languages, and exporting outputs. Additionally, the system will now focus on enhancing accessibility, real-time performance, and integrating ethical AI practices.

1.2 Scope

Textify shall provide real-time text translation, text recognition from images via OCR, and support for exporting translations in multiple formats. The application will ensure interoperability through API integrations and a user-friendly interface accessible via web browsers on both desktop and mobile platforms. The updated focus emphasizes adaptive learning, multilingual support for the UI, and compliance with data privacy regulations.

1.3 Definitions, Acronyms, and Abbreviations

- **API:** Application Programming Interface
- **OCR:** Optical Character Recognition
- **NMT:** Neural Machine Translation
- **TTS:** Text-to-Speech
- **STT:** Speech-to-Text
- **GDPR:** General Data Protection Regulation
- **Accessibility:** Compliance with standards like WCAG 2.1

1.4 Overview

This document describes the functional and non-functional requirements for Textify, with a focus on AI-enhanced features such as OCR, TTS, and STT, ethical AI practices, and user-centered design. It includes updated roles and responsibilities, a revised system architecture, and use cases that align with current advancements in AI-enabled software engineering.

2. General System

2.1 Product Overview

- Textify shall provide accurate and efficient text translation, OCR, TTS, and STT functionalities.
- The system will integrate advanced AI models tailored for domain-specific accuracy and performance.

- Real-time processing will be prioritized to enhance user experience.
- The application will be web-based and designed for accessibility across standard web browsers on desktop and mobile platforms.

2.2 Roles and Responsibilities

- **Customer:** Defines high-level requirements and reviews progress.
- **Project Manager:** Oversees timelines, allocates resources, and manages deliverables.
- **AI Specialist:** Fine-tunes and integrates AI models, ensuring high performance and accuracy.
- **Programmers:** Implement core features, including OCR, translation, and TTS/STT.
- **DevOps Team:** Ensures scalability, deployment, and performance optimization.
- **Testers:** Validate features against acceptance criteria, focusing on edge cases and accessibility compliance.

3. Functional Requirements

3.1 Text Translation

- **Requirement:** Accept input text in a source language and translate it to a target language with real-time feedback.
- **Acceptance Criteria:** Translations for text under 500 words shall be returned within 1.5 seconds.

3.2 Text-to-Speech (TTS)

- **Requirement:** Convert text to speech with customizable playback controls (pause, stop, resume) and export functionality.
- **Acceptance Criteria:**
 1. Provide options to adjust speed and volume.
 2. Support download in standard audio formats.

3.3 Speech-to-Text (STT)

- **Requirement:** Transcribe spoken input to text with 90% accuracy in moderately noisy environments.
- **Acceptance Criteria:** Include punctuation and capitalization in transcriptions.

3.4 Optical Character Recognition (OCR)

- **Requirement:** Extract and translate text from images with 95% accuracy for standard fonts and adaptive preprocessing.
- **Acceptance Criteria:**
 1. Accept PNG, JPEG, and PDF inputs.
 2. Preprocess images (e.g., de-skewing, noise reduction) before extraction.

3.5 User Interface

- **Requirement:** The application shall have a responsive, multilingual design compliant with WCAG 2.1.
- **Acceptance Criteria:**
 1. Key features load within 1 second on standard devices.
 2. Offer multilingual support for UI elements.

3.6 Exports

- **Requirement:** Enable users to export translations into formats such as PDF, Word, or TXT.
- **Acceptance Criteria:** Provide a clear, intuitive interface for export options.

3.7 API Integration

- **Requirement:** Allow integration with third-party applications to enhance functionality.
- **Acceptance Criteria:** Provide well-documented REST APIs for translation and OCR services.

4. Non-Functional Requirements

4.1 Scalability

- **Requirement:** Support up to 10,000 concurrent users with a response time under 3 seconds.

4.2 Security

- **Requirement:** Encrypt all data transmissions using SSL/TLS. Ensure compliance with GDPR.

4.3 Privacy

- **Requirement:** Obtain explicit consent before using user data for model training.

4.4 Maintainability

- **Requirement:** Follow a modular architecture with centralized dependency management.
- **Acceptance Criteria:**
 1. Use inline comments and detailed documentation.
 2. Employ version control for code management.

4.5 Reliability

- **Requirement:** Provide error detection and logging with clear error messages for debugging.
- **Acceptance Criteria:** Ensure logs include timestamps and error codes.

4.6 Accessibility

- **Requirement:** Ensure compliance with WCAG 2.1 standards for all UI components.

5. Use Cases

5.1 Text Translation

- **Actors:** User
- **Description:** Translate input text between languages.
- **Postcondition:** Display translated text.

5.2 OCR (Image-to-Text)

- **Actors:** User
- **Description:** Upload an image for text extraction.
- **Postcondition:** Display extracted text.

5.3 Export Translation

- **Actors:** User
- **Description:** Export translations in various formats.
- **Postcondition:** Download completed.

5.4 STT

- **Actors:** User
- **Description:** Transcribe uploaded or live audio input.
- **Postcondition:** Provide transcribed text.

5.5 TTS

- **Actors:** User
- **Description:** Convert input text into an audio file.
- **Postcondition:** Provide audio file.

6. System Design and Architecture

6.1 Architecture Overview

- **Frontend:** Responsive web-based interface.
- **Backend:** AI-enhanced services for OCR, translation, TTS, and STT.
- **Database:** Secure storage for user data, logs, and translation history.

6.2 Revised Features

- Incorporate adaptive AI learning for better accuracy.
- Ensure interoperability through REST APIs.
- Prioritize accessibility and multilingual support in UI design.