Requirements Specification for Textify

—

General System Requirements

Category	Requirement
Product Overview	Textify supports translation, OCR, TTS, STT, and file export options. Web-based, user-friendly interface.
Roles & Responsibilities	Roles include Customer, Manager, Designer, Programmers, Testers.

. Functional Requirements

Feature	Requirement	Acceptance Criteria
Text Translation	Translate text between source and target languages.	User inputs text and selects languages; translation is returned.
TTS (Text-to-Speech)	Convert text (including translations) to audio output.	Text-to-speech conversion works; users can control rate and volume.
STT (Speech-to-Text)	Accurately transcribe speech into written text.	Audio is transcribed into text with punctuation and capitalization.
OCR (Optical Character Recognition)	Extract text from images (e.g., scanned documents, photos with text).	Text from images is correctly extracted and displayed; preprocessing is done.
Export	Support export of translations, OCR text, and transcribed speech in multiple formats (PDF, TXT, DOCX).	User can export translations and text in selected formats.

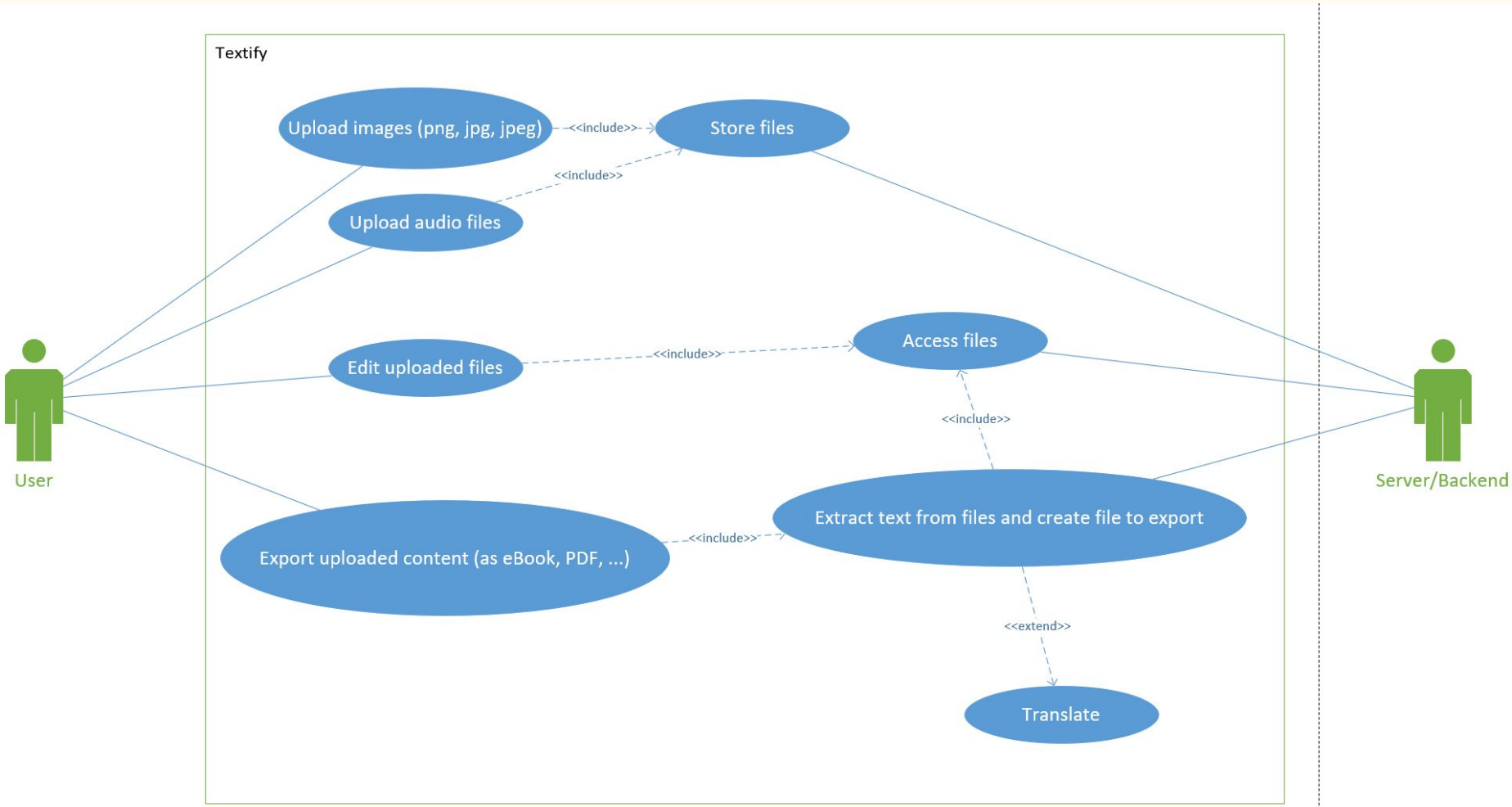
Non-Functional Requirements

Category	Requirement	Acceptance Criteria
Scalability	System must be able to handle a growing number of requests and users, particularly during peak demand periods.	The system should support increased user load without performance degradation.
Security	Data transmissions must be encrypted using SSL/TLS.	All communications must be encrypted and user data should not be stored unencrypted.
Privacy	User data should not be used for model training unless explicit consent is obtained.	User data is not used for training without explicit user consent.
Maintainability	The system must have a modular architecture to allow for easy updates, bug fixes, and future feature additions.	System's architecture should be modular, enabling easy updates and isolated component changes.
Automated Testing	The system should have a comprehensive suite of automated tests to verify correctness and ensure new bugs are not introduced.	Unit tests should verify the core system functions (translation, OCR, TTS, STT, export).
Error Logging	A robust error logging mechanism must be in place to detect issues in real time and support troubleshooting.	The system logs errors, warnings, and critical events with enough detail for troubleshooting.
Version Control	Use of version control (e.g., Git) for tracking code changes and collaboration among developers.	All code changes should be tracked using Git with clear commit messages.
Dependency Management	The system must manage external libraries and dependencies effectively.	Dependencies should be clearly listed, versioned, and managed (e.g., package.json, requirements.txt).

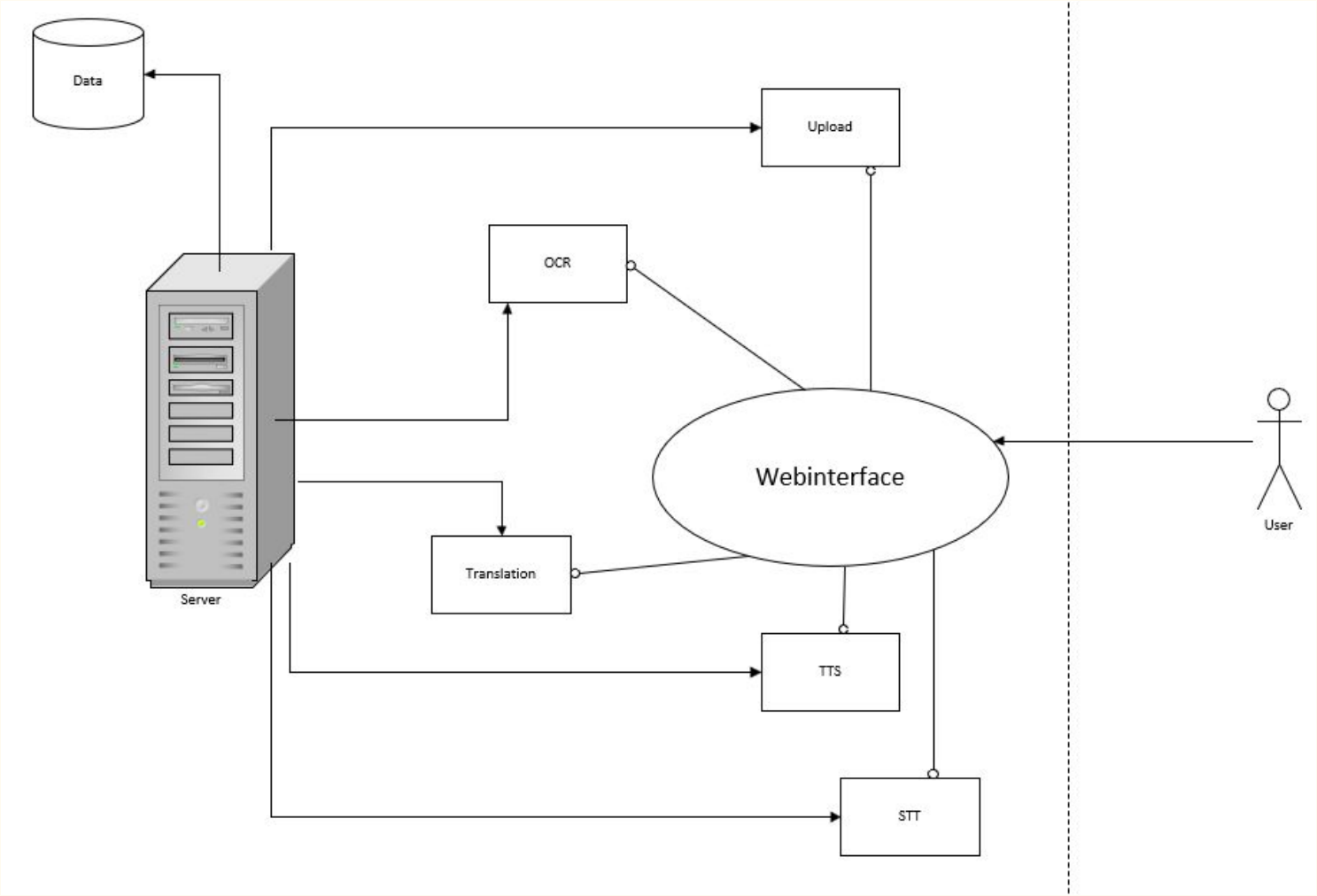
Easy Configuration and Deployment	The system should allow easy configuration and deployment to staging, testing, and production environments.	Configuration files and deployment scripts should allow easy setup and consistent environments.
Image Preprocessing	The system must preprocess images (e.g., de-skewing, noise reduction) to improve OCR accuracy.	Preprocessing techniques should enhance OCR text extraction quality.
Error Handling	The system must provide user-friendly error messages and prevent errors before they occur.	Users should receive clear, actionable error messages and input validations (e.g., file formats, text length).
User Experience	The system must be responsive and optimized for various devices (desktop, mobile).	The interface should be dynamic and adapt to different screen sizes without horizontal scrolling.
Multilingual Support	OCR and TTS should support multiple languages.	OCR and TTS must be able to recognize and process text and speech in various languages, including non-Latin scripts.

Use Cases Summary

Use Case	Actors	Description	Preconditions	Postconditions
Text Translation	End User	A user inputs text to be translated into another language.	The user has launched the system and is on the translation page.	The system returns a translation of the input text.
OCR (Image-to-Text)	End User	A user uploads an image, and the system extracts and displays text from the image.	The user has uploaded an image containing text (e.g., scanned document or photo).	The system extracts and displays the text from the image.
Speech-to-Text (STT)	End User	A user uploads an audio file which is transcribed into text.	The user has uploaded an audio file.	The system transcribes the speech into text.
Text-to-Speech (TTS)	End User	A user inputs text, and the system converts it to speech.	The user has entered text for conversion into speech.	The system generates an audio file of the text and allows playback or download.
Export Translations	End User	A user exports translated text to a different format (e.g., PDF, Word, TXT).	The user has received a translation.	The user can download the translated text in the selected format (e.g., PDF, TXT).


Textify
Upload images (png, jpg, jpeg)
<<include>>
Store files
<<include>>
Upload audio files
Edit uploaded files
<<include>>
Access files
<<include>>
Export uploaded content (as eBook, PDF, ...)
<<include>>
Extract text from files and create file to export
<<extend>>
Translate
User
Server/Backend

Component	Description
Frontend	Web-based user interface for interacting with the system.
Backend	Server hosting the AI models for translation, speech synthesis, and OCR processing, as well as APIs for communication.


Data
Upload
OCR
Server
Webinterface
User
Translation
TTS
STT