

Software Requirements Specification Document (SRS)

1. Introduction

1.1 Purpose

The purpose of this document is to describe the functional and non-functional requirements of the Reel2MovieName system. This system identifies the original movie from short video clips using multimodal analysis including audio, visual frames, and on-screen text.

1.2 Scope

The system allows users to upload short video clips obtained from platforms such as Instagram Reels or YouTube Shorts. The system processes the clip using audio recognition, visual similarity detection, and text extraction techniques to identify the corresponding movie. After identification, detailed movie information is retrieved from a movie database API and displayed to the user.

1.3 Definitions, Acronyms, and Abbreviations

FR – Functional Requirement

NFR – Non-Functional Requirement

UI – User Interface

DB – Database

OCR – Optical Character Recognition

API – Application Programming Interface

1.4 References

[1] IEEE 830-1998 Standard for Software Requirements Specification

[2] TMDB API Documentation

[3] ACRCLOUD Audio Recognition API

[4] Sommerville, Software Engineering

2. Overall Description

2.1 Product Perspective

Reel2MovieName is a web-based application integrated with AI-based video analysis modules and external movie database APIs. The system processes short video clips temporarily and does not store copyrighted content permanently.

2.2 Product Functions

Main functions of the system include:

- User Registration and Login
- Video Upload
- Audio Recognition
- Frame Extraction
- Text Extraction (OCR)
- Movie Identification
- Display Movie Information

2.3 User Classes

- User
- Admin

2.4 Operating Environment

Hardware Requirements:

- Processor: Intel i5 or higher
- RAM: Minimum 8 GB
- Storage: 10 GB free space
- Camera/Video-enabled device

Software Requirements:

- OS: Windows 10/11 or Linux
- Browser: Chrome, Edge, Firefox
- Backend: Python (FastAPI)
- Database: MySQL / SQLite
- APIs: TMDB API, ACRCLOUD API

3. Specific Requirements

3.1 Functional Requirements

FR1: User Registration

The system shall allow users to register using valid credentials.

FR2: User Login

The system shall authenticate users using username and password.

FR3: Video Upload

The system shall allow users to upload short video clips.

FR4: Audio Extraction

The system shall extract audio from the uploaded video.

FR5: Audio Recognition

The system shall identify movie audio using ACRCLOUD API.

FR6: Frame Extraction

The system shall extract frames from the video.

FR7: Visual Similarity Detection

The system shall compare frames using AI models (CLIP).

FR8: Text Extraction

The system shall extract on-screen text using OCR.

FR9: Multimodal Decision

The system shall combine audio, visual, and text analysis results to identify the movie.

FR10: Movie Data Retrieval

The system shall retrieve movie details from TMDB API.

FR11: Result Display

The system shall display movie name, poster, cast, and description.

FR12: Error Handling

The system shall notify the user when movie identification fails.

FR13: History

The system shall store search history for logged-in users.

FR14: Logout

The system shall allow users to logout securely.

3.2 Non-Functional Requirements

NFR1: Performance

The system shall process a video and return results within a few seconds.

NFR2: Scalability

The system shall support multiple concurrent users.

NFR3: Security

The system shall encrypt user data and prevent unauthorized access.

NFR4: Reliability

The system shall provide accurate and consistent movie identification.

NFR5: Availability

The system shall be available 24/7.

NFR6: Usability

The system shall provide a simple and user-friendly interface.

NFR7: Maintainability

The system shall be easy to update and maintain.

NFR8: Compatibility

The system shall work on different browsers and devices.

NFR9: Data Integrity

The system shall prevent unauthorized modification of stored data.

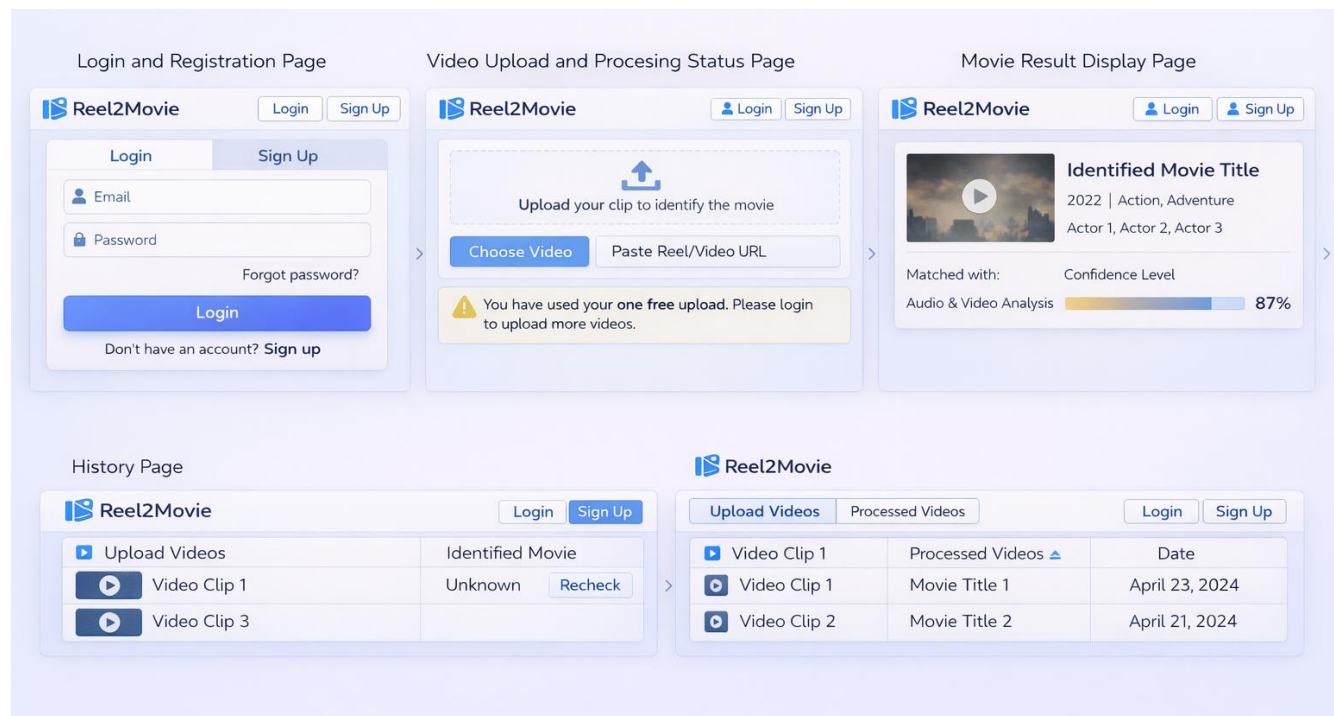
NFR10: Backup & Recovery

The system shall support data backup and recovery.

4. External Interface Requirements

4.1 User Interfaces

- Login and Registration Page
- Video Upload Page
- Processing Status Page
- Movie Result Display Page
- History Page



4.2 Hardware Interfaces

The system requires a video-enabled device. No special hardware is required.

4.3 Software Interfaces

System	Purpose	Data Format
Frontend (React.js)	User Interface	JSON
Backend (FastAPI)	Business Logic	REST/JSON
TMDB API	Movie Information	JSON
ACRCloud API	Audio Recognition	JSON
OCR Engine	Text Detection	Image/Text

4.4 Communication Interfaces

- HTTP/HTTPS
- REST API
- JSON Data Format

5. System Features

5.1 Use Case Diagram

Actors

- User
- Admin

Use Cases

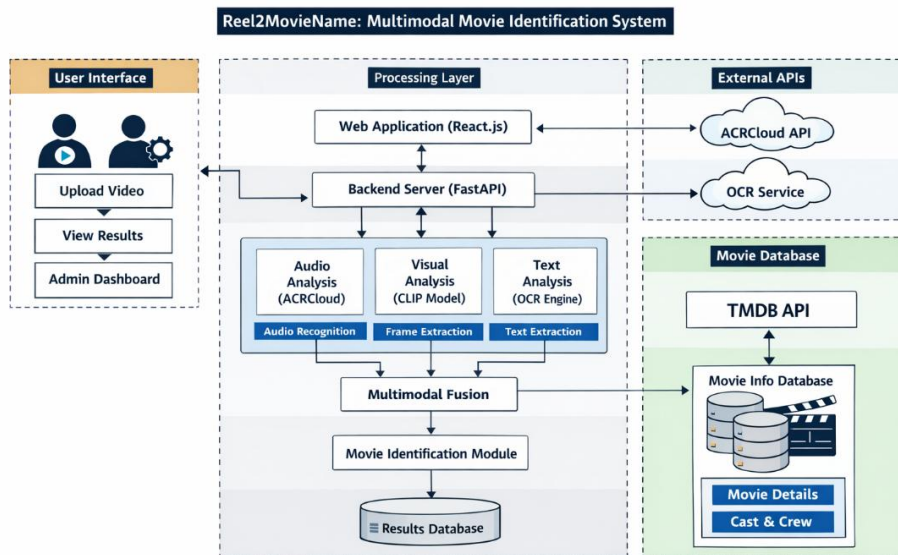
User:

- Register/Login
- Upload Video
- View Movie Details

Admin:

- Login

- Monitor System
- Manage Users



6. Data Requirements

6.1 Database Design

- Users
- Videos
- Results
- History

6.2 Table Structures

1. Users Table

Field Name	Data Type
user_id (PK)	INT
name	VARCHAR
email	VARCHAR (100)
password	VARCHAR (50)

created_at	DATETIME
------------	----------

2. Videos Table

Field Name	Data Type
video_id(PK)	INT
user_id(FK)	INT
file_path	VARCHAR (100)
upload_time	DATETIME

3. Results Table

Field Name	Data Type
result_id(PK)	INT
video_id(FK)	INT
movie_name	VARCHAR (100)
confidence_score	FLOAT

6.3 Data Retention

Video files are processed temporarily and deleted after result generation. User data is retained unless deletion is requested.

7. Quality Attributes

7.1 Availability

System shall be available 24/7 except scheduled maintenance.

7.2 Security

System shall use authentication, authorization, and encrypted communication.

7.3 Performance

Response time shall be optimized using efficient AI models and APIs.

7.4 Maintainability

System shall be modular and well-documented.

8. Glossary

OCR – Text extraction from images

API – Interface to communicate between systems

CLIP – Vision-language AI model

Multimodal – Use of multiple data sources

References

- [1] IEEE. (1998). *IEEE 830-1998 Standard for Software Requirements Specification*. IEEE Computer Society.
- [2] Radford, A., et al. (2021). *Learning Transferable Visual Models From Natural Language Supervision (CLIP)*. OpenAI.
<https://openai.com/research/clip>
- [3] ACRCLOUD. (2024). *Audio Recognition API Documentation*.
<https://www.acrccloud.com/docs/>
- [4] The Movie Database (TMDb). (2024). *TMDb API Documentation*.
<https://developer.themoviedb.org/>
- [5] Smith, R. (2007). *An Overview of the Tesseract OCR Engine*. Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR).
- [6] Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed.). Pearson Education.
- [7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [8] FastAPI Documentation. (2024).
<https://fastapi.tiangolo.com/>
- [9] React.js Documentation. (2024).
<https://react.dev/>
- [10] Sommerville, I. (2015). *Software Engineering* (10th ed.). Pearson Education.