

FLORIDA INTERNATIONAL UNIVERSITY



VIRS – Vocabulary in Reading Study

Product Owner: Dr. Eric Dwyer

Professor: Masoud Sadjadi

Team: George Spanioli, Issik Drew, Joshua Lopez, Ivan Marquez Jr, Justin Robertson

What is VIRS?

VIRS or vocabulary in reading study is a project that aims to assist students learning english, and teachers trying find out which vocabulary to teach them. If the most used vocabulary is learned, it is possible to learn a new language at a faster rate.

Word List

- K1- Among the list of the 1000 most frequently used words in primary and secondary texts.
- K2- Among the list of the 2nd 1000 most frequently used words in primary and secondary texts.
- K3- Among the list of the 3rd 1000 most frequently used words in primary and secondary texts.
- Other High Frequency
- Medium Frequency

Word List

- Low Frequency
- Academic Words- Commonly occurring among a wide variety of academic subjects but not within the 2000 most frequent words.
- STEM Words- Words occurring in Math or Science texts but not within the 2000 most frequent words.
- Names & Off-List

What are the site's features?

- Enhance text to show words highlighted to the color of the category from the frequently used words.
- Get statistics of vocabulary words on typed text, word docs, PDF, and images to see what level of reading it is and how many frequently used words there are.
- Able to translate typed text to up to 60+ languages
- Download the lists of vocabulary words.

Problems

1. At the beginning, the website was at a non-operational state.
2. The UI of the site was outdated.
3. Limit of up to 25 MB when uploading files.
4. Uploading files would sometimes not work or just wouldn't work with the “Enhance text” feature.

Object Design

1. Text Processor: Parses uploaded text, tokenizes words, and cleans punctuation.
2. Vocabulary Analyzer: Calculates word frequency and difficulty using corpus statistics and identifies key academic terms.
3. User Interface Controller: Manages pages for uploading texts and displaying analysis results to teachers.
4. Database Manager: Handles connections to the word-frequency database and caches analysis results.

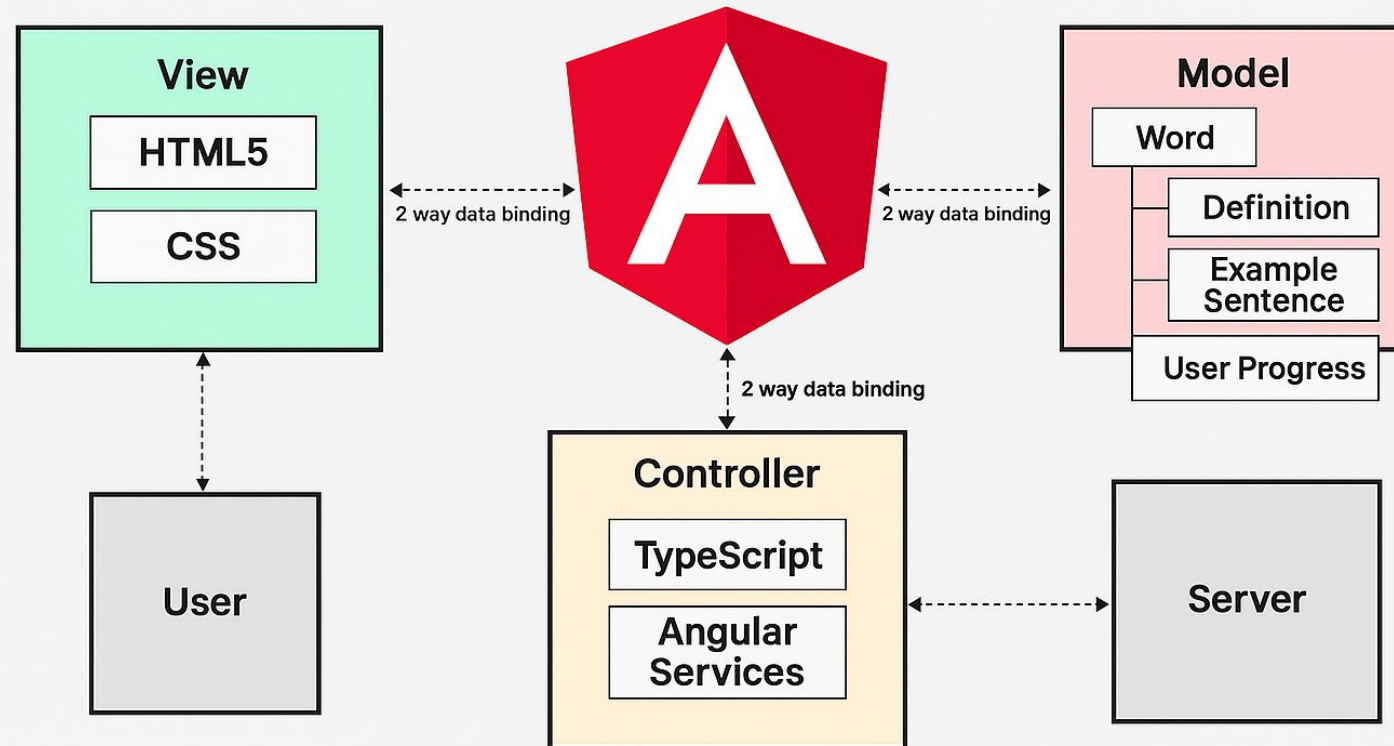
System Design

The system described in the image outlines a full-stack web application implementation. The frontend is developed using Angular and Bootstrap, which ensures that the interface is responsive, modern, and user-friendly across devices.

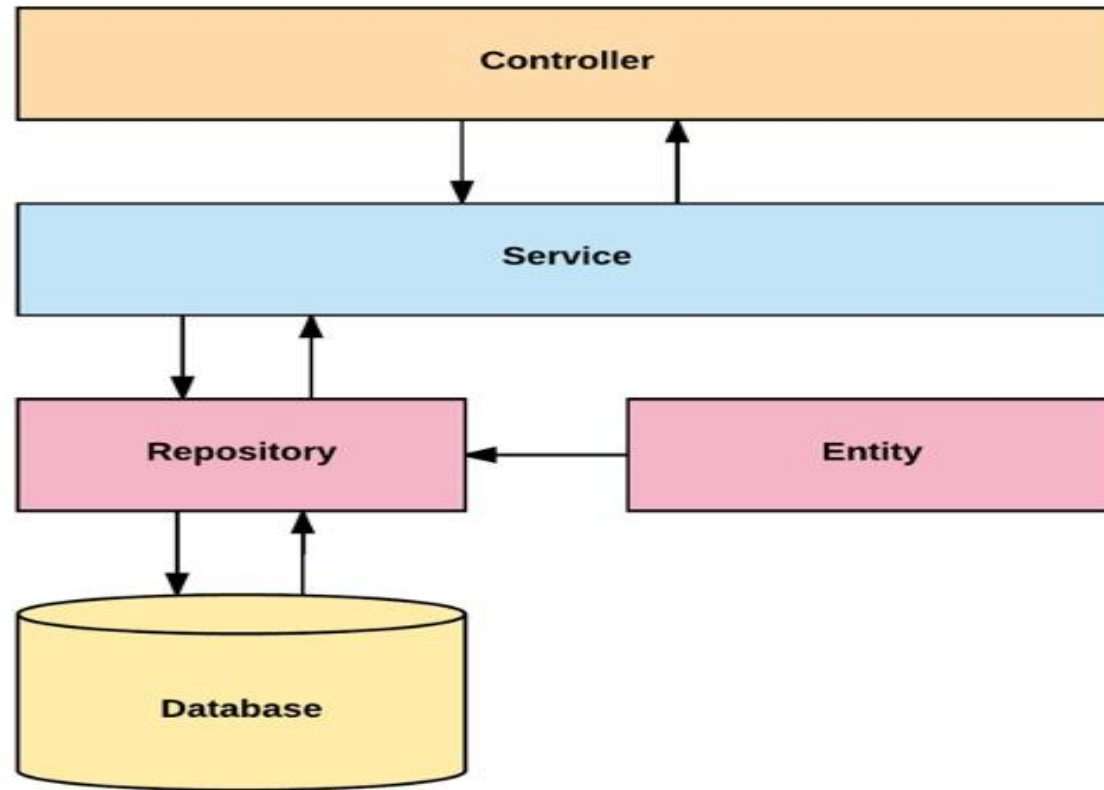
The backend leverages Java Spring Boot for web service architecture and integrates Python NLP libraries like spaCy and NLTK to perform natural language processing tasks, such as text parsing and vocabulary analysis.

The application stores word frequency and difficulty data in a MySQL database, which the backend accesses and serves to the frontend via RESTful APIs. Finally, the entire application is hosted and deployed on AWS Elastic Beanstalk, a cloud platform that provides scalability and reliability for managing and running web applications efficiently.

Front end



Back end



User Stories



- **User Story #1** — User Authentication System: As a user, I want to sign up and log in securely, so my progress is saved and private. (Frontend + Backend)
- **User Story #2** — Reading Interface: As a user, I want to read text passages with vocabulary words highlighted so I can focus on learning those words. (Frontend)
- **User Story #3** — Vocabulary Management: As an admin, I want to upload reading passages and related vocabulary so that learners always have fresh content. (Backend + Database)
- **User Story #4** — Word Lookup Feature: As a user, I want to click a highlighted word to see its definition and hear it pronounced. (Frontend + Backend)
- **User Story #5** — Flashcard Review Game: As a user, I want to practice words using flashcards that show the definition and optionally play the pronunciation. (Frontend)

User Stories



- **User Story #6** — Progress Tracking System: As a user, I want to see how many words I've practiced and mastered so I can track my learning. (Frontend + Backend + Database)
- **User Story #7** — Mobile-Friendly Design: As a user, I want the app to work well on phones and tablets so I can use it anywhere. (Frontend + QA)
- **User Story #8** — User Feedback Submission: As a user, I want to leave feedback on passages or vocabulary content so that the app can improve based on user suggestions. (Frontend + Backend)
- **User Story #9** — Session Stability Fixes: Addressed random logout issues by patching session management on the backend.
- **User Story #10** — Admin Dashboard Analytics (Finalized): Backend data queries were optimized, and full engagement statistics are now available to admins.

User Stories



- **User Story #11** – Mobile UI Enhancements: Refined navigation for mobile users, including menu responsiveness and component spacing. (Frontend only)
- **User Story #11** – Vocabulary Word Reporting Tool: As a user, I want to report incorrect definitions or pronunciations so that content can be corrected. (Frontend + Backend + QA)
- **User Story #14** – Bookmark Favorite Words: As a user, I want to save and revisit specific vocabulary words across all devices. (Frontend + Backend)



Summary

Our system ingests classroom texts, analyzes vocabulary, and highlights words by frequency and difficulty, empowering teachers to deliver targeted instruction. We implemented the tool using Angular, Spring Boot, and Python NLP libraries, achieving accurate and scalable results. Ongoing validation and teacher feedback confirm its usefulness.