Department of Computer Science, University of Hong Kong, COMP4801 Final Year Project Interim Report

All-In-One Dyslexic Learning Assistant Platform

Supervisor:

Prof. Luo Ping

Second Examiner:

Prof. Liang Ying Yu

Author:

Chung Cheuk Man (3035928079)

Sin Chung Hang (3035928067)

Tiquia Joezen S. (3035930096)

Lo Lok Fung (3035930228)

Date: January 20, 2025

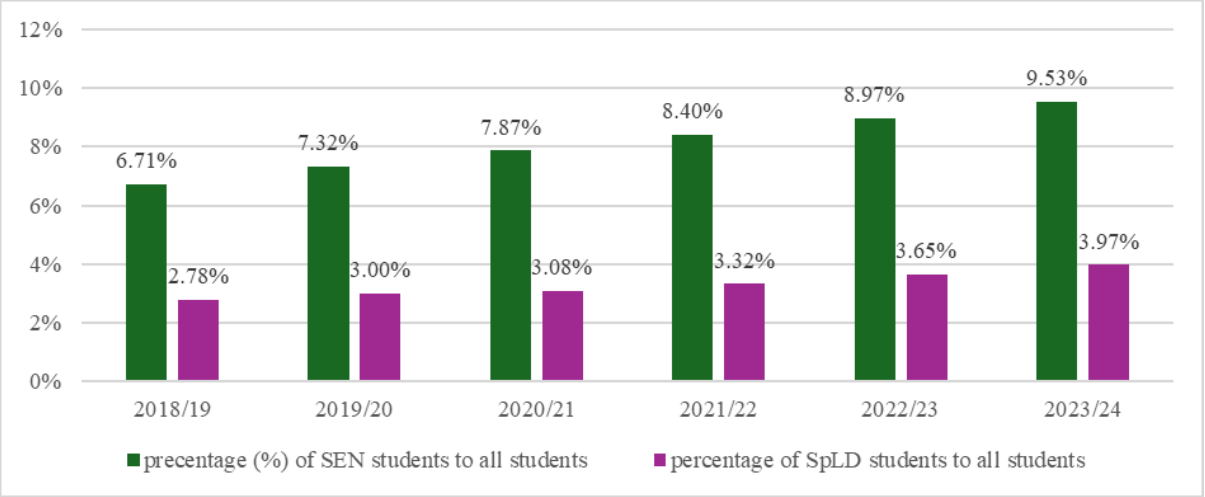
1. Introduction

Hong Kong's education system is renowned for its competitiveness and high academic standards. However, the system faces an underlying issue and challenge: the needs of Special Educational Needs (SEN) students. Focusing solely on primary students, there are approximately 300,000 (Education Bureau, 2024a). Of these, around 10 percent are identified as SEN students, which is concerning since those account for 30,000 students (Education Bureau, 2024b). The Education Bureau (2021) guidelines state that nine different types of SEN can receive support from the government.

The nine types are as follows (Education Bureau, 2021):

1. Mental Illness
2. Physical Disability
3. Hearing Impairment
4. Speech and Language Impairment
5. Visual Impairment
6. Autism Spectrum Disorder
7. Intellectual Disability
8. Attention Deficit / Hyperactivity Disorder
9. Specific Learning Difficulties

The motivation for the project is that a majority of SENs have dyslexia. As stated in a paper published by the Education Bureau in 2021, the most common type of SEN among primary students is specific learning difficulties (SpLD), which include dyslexia.



According to the International Dyslexia Association (2018), dyslexia is a specific learning disability that is neurobiological in origin. It is marked by challenges in accurate and fluent word recognition, as well as difficulties in decoding and spelling. These difficulties typically result from a deficit in the phonological component of language that is often unexpected concerning other cognitive abilities and the provision of effective classroom instruction. Secondary effects can include difficulties in reading comprehension and a diminished reading experience, which may hinder the development of vocabulary and background knowledge.

This project aims to help dyslexic students overcome all the challenges by developing an all-in-one dyslexic learning assistant platform. The main audience will be kindergarten to primary students, and the platform will mainly help with Chinese dyslexia.

2. Objectives

The platform will be developed by focusing on four key objectives. There are four subsections: Section 2.1: Auditory-guided visualization; Section 2.2: Converting complex phrases to simple language; Section 2.3: Learning through voice communication; and Section 2.4: Learning through gaming.

2.1. Auditory-guided visualization

Auditory-guided visualization enhances learning by using auditory stimuli to help students create mental images and improve comprehension. This approach is particularly beneficial for students with dyslexia, who may struggle with traditional reading methods.

2.2. Converting complex phrases to simple language

The primary aim of converting complex phrases into simple language is to make educational materials more understandable and accessible, especially for students with dyslexia.

2.3. Learning through voice communication

Integrating voice into learning fosters more engagement and understanding for students, including those with learning differences such as dyslexia. Speech recognition tools will enable students to dictate their thoughts verbally, which could be transcribed into written text. This would eliminate the anxiety of writing experienced by students who struggle to put their thoughts onto paper due to dyslexia.

2.4. Learning through gaming

Learning through gaming uses interactive and engaging games to make education interesting and effective. Games for dyslexic students help reinforce their skills in multisensory ways while boosting their motivation and confidence within a fun, low-pressure environment.

3. Methodology

3.1. Main Tech Stack

3.1.1. Frontend (React.js)

Regarding the front end of the project system, React.js is adapted as the front-end framework. React.js is a popular front-end development framework worldwide. Therefore, it has a rich ecosystem that provides a wide variety of libraries or packages for front-end development. Also, its component-based architecture allows the project to be expanded or scaled easily, as the front-end elements can be divided into components for development, which will be reusable in the future.

To further enhance the frontend development, the Tailwind CSS framework is used to apply styles in HTML directly for fast styling. The React Vite building tool also provides a lightning-fast frontend development server with hot module replacement.

3.1.2. Backend (Python Flask)

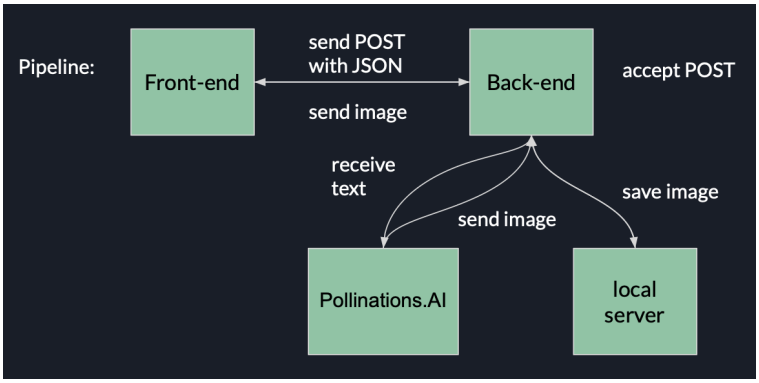
Python Flask is chosen as the backend framework for the project's backend field. It is a relatively lightweight and flexible framework that helps to build the application without unnecessary overheads. Also, the backend of the application could be built quickly due to its simplicity. Besides, Python could easily be used to integrate with third-party services like Azure or online API services, which is suitable for our projects.

3.2. Vocabulary/phrases-to-image translation

3.2.1. Overview

In this feature, the student could input a simple word or vocabulary to generate an image corresponding to this vocabulary. Therefore, the dyslexic student could gain a more thorough understanding of the word by the image.

Pollination AI will be used as the vocabulary-to-image tool. It is an open-source AI model to generates high-quality images based on the user prompt.



The overall pipeline of the feature is shown in the above flowchart. The user could send a post request with the input vocabulary as the body to the backend server. Then, the backend server will communicate with the pollination AI to get the image that corresponds to the vocabulary. The image will be saved on the local server. After that, the image will be sent back to the user front so that the user can observe the image corresponding to the vocabulary.

3.2.2. Frontend

Regarding the frontend feature, the vocabulary-to-image feature will be a particular tab in the system. Inside this tab is a vocabulary input field for the students to enter the vocabulary they would like to search for. Below the vocabulary input field, there is an image-showing field to display the image corresponding to the vocabulary. Both Chinese and English vocabulary could be accepted in the vocabulary input field.

3.2.3. Backend

The backend server consists of two major functions:

1. Download image function

   It sends a GET request to the pollination API to fetch the image corresponding to the word. The image is then saved locally, and this function returns it to the frontend server.

2. Generate image function

   It receives the JSON data sent from the front end and retrieves the vocabulary inside the post request. Then, it calls the download image function to communicate with the pollination AI.

3.2.4. Limitation

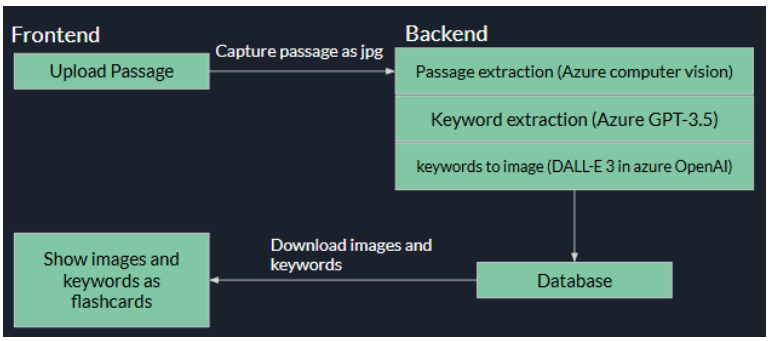
However, there is a limitation when developing this feature. The response time from pollination.ai will be too long due to the long input text.

There are two possible solutions to resolve the above situation. First, some online tools can be integrated to capture the key text of the input field. Therefore, the input text sent in the post request could be shortened to reduce the response time of the text-to-image function. Secondly, the input field could be limited by its input length to avoid the user sending long text to the pollination AI.

3.3. Short text revision flashcard generator

3.3.1. Overview

This project aims to create a flashcard generation tool designed to aid SEN students in recognizing and understanding short text passages. The system accepts a passage in JPEG format and produces flashcards displaying relevant keywords alongside corresponding images. This approach leverages visual learning, enhancing comprehension and retention for students who benefit from image-based learning materials.



The JPEG passage is passed to Azure Computer Vision for recognizing the passage outputted as a text string, which is passed to the keyword extraction function to extract keywords using Azure GPT-3.5 (with the number of keywords depending on the prompt). Each keyword is drawn to generate the corresponding image, and the keyword-to-image mapping is then into the database, which will be extracted by the front end to make flashcards with the corresponding keywords and images.

The system architecture is divided into frontend and backend components, communicating through a database.

3.3.2. Frontend

The frontend, built with ReactJS, employs two main components: *Flashcard* and *FlashcardArray*. The *Flashcard* component renders individual flashcards, each presenting a keyword and its associated image. The *FlashcardArray* component fetches keyword-image pairs from the backend database and dynamically renders a collection of *Flashcard* components, creating an interactive flashcard deck for the user.

3.3.3. Backend

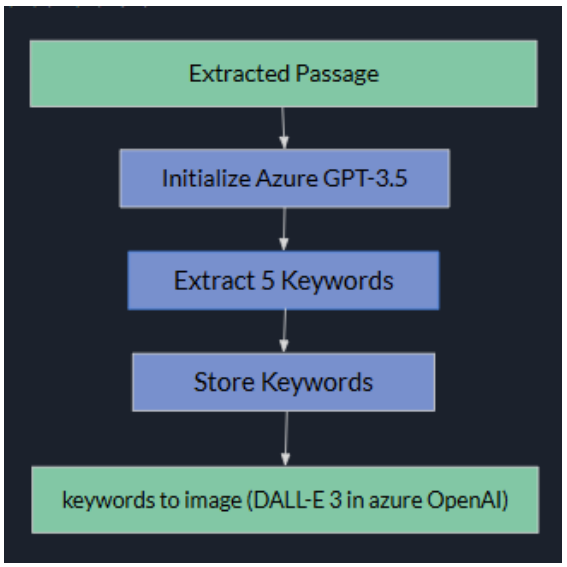
The backend leverages the power and scalability of Azure Cloud Services. Specifically, it utilizes Azure Computer Vision, Azure GPT-3.5, and DALL-E 3. These services were

chosen for their cost-effectiveness, reliability, scalability, and ease of management, crucial factors for a project in its initial development phase.

3.3.3.1. Input Processing

The process begins with optical character recognition (OCR) using Azure Computer Vision.

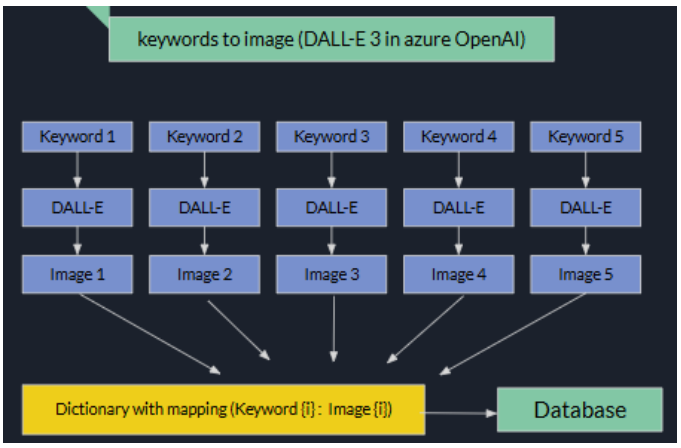
The process is shown below in Fig.



This service accurately converts the input JPEG image containing the text passage into a text string. This extracted text is then fed to Azure GPT-3.5 for keyword extraction. A carefully crafted prompt guides GPT-3.5 to identify a specified number of relevant keywords, ensuring they are non-violent, distinct, and feature complex character strokes suitable for SEN student recognition.

3.3.3.2. Image Generation

The pipeline of the image generation with the utilization of DALL-E 3 is shown in the figure below.



Once keywords are extracted, the system iteratively generates corresponding images. Each keyword is passed to DALL-E 3, which produces a representative image. This keyword-image pair is then stored in the database. This process repeats until images have been generated for all extracted keywords. The database then serves as the source for the frontend, providing the necessary data to create the flashcards.

3.4. Voice Chatbot

The Voice Chatbot has a well-integrated voice input/output system for speech-to-text and text-to-speech interactions, enabling users to speak their will for an advanced user experience. It offers multi-language support: Cantonese, Putonghua, and English, allowing users to practice the different languages learned in school. Also, under the strong backup of Azure OpenAI, it is capable of discussing a specific passage with the user and raising his or her comprehension while nudging toward interactive learning.

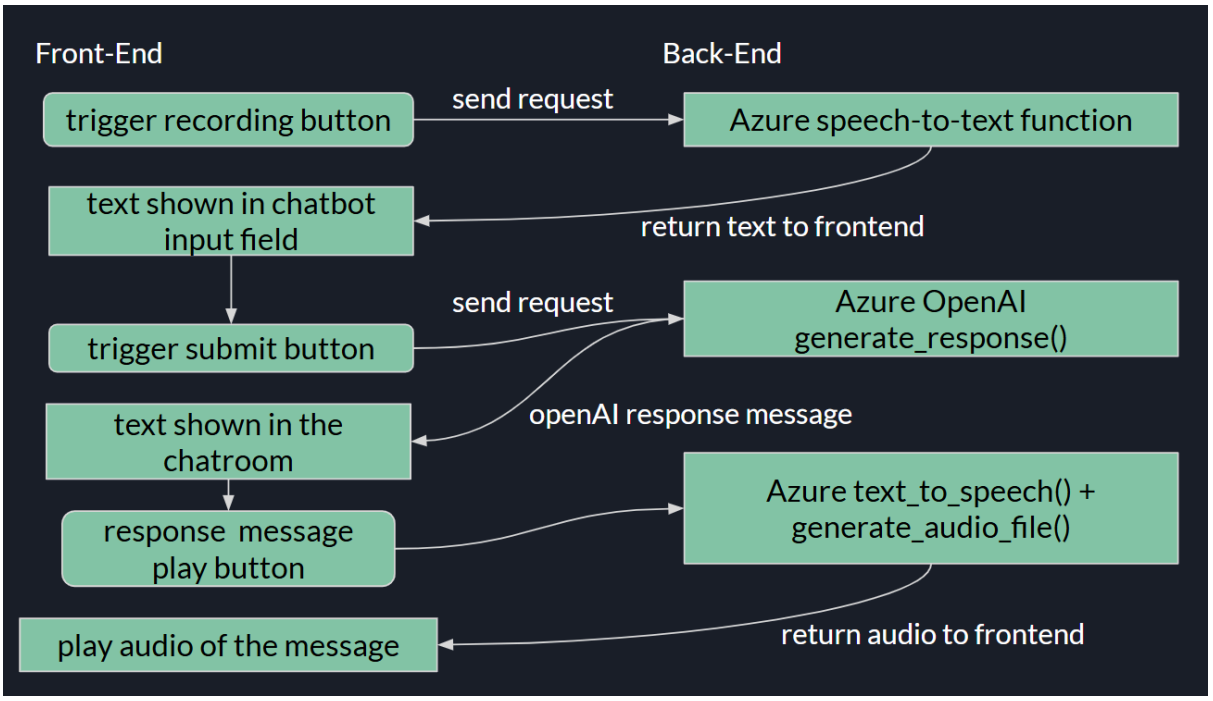
3.4.1. Frontend

The chatbot is built with ReactJS.

3.4.2. Backend

The backend is mainly developed in Python. The libraries used are Azure.cognitiveservices.speech, basic conversation handling, and Azure OpenAI. The backend utilizes Azure Services, specifically Azure Speech Services, for professional speech recognition and synthesis. This allows the chatbot to recognize speech input and synthesize speech output in Cantonese, Putonghua, and English with outstanding clarity, supported by different voice models for each language. By doing so, the robust framework elevates user interaction and ensures access to diverse linguistic needs.

3.4.3. Workflow



From Figure xx, the chatbot works by recording speech by clicking the audio record button. This sends a request to the Azure speech-to-text function in the backend to convert the recorded audio. The text is returned to the frontend and shown on the chatbot input field. It is optional for the user to edit the text by typing. By triggering the submit button, a request is

sent to the Azure OpenAI API to generate a response from the AI. In the chatroom on the frontend, the response message will be displayed. Users can listen to the audio response by clicking the play button below the response message. This requests the backend to utilize the Azure text-to-speech function to generate an audio file. It will then be played on the frontend.

3.5. Gamified Exercise

Playing games allows kids to learn through play, which is again very effective and entertaining, especially for dyslexic children. Games can make learning more engaging and fun and enhance abilities like critical thinking and problem-solving. Our dyslexia support system will include three different game categories: character puzzle, character bingo, and flashcard frenzy. This strategy ensures an enjoyable and encouraging learning environment catered to their needs.

3.5.1. Game #1: Character Puzzle

Using their radicals to understand Chinese characters on a structural level, the Character Puzzle game helps dyslexic students develop their orthographic skills. In this game, the user matches the target character by arranging the provided radicals.

To make the target character's proper structure, players drag them into the designated location below. To encourage perseverance and reinforce learning via trial and error, the game provides feedback on whether the user has successfully designed the character or if it needs revisions.

Interacting with radicals helps one understand how characters are constructed, which is essential for reading and writing proficiency. The development of visual memory and recognition aids the dyslexic student in overcoming obstacles related to character development. By doing this, one gains confidence in their ability to understand a potentially

challenging character visually and in its production as they proceed through increasingly higher levels, resulting in an enhanced overall literacy outcome.

3.5.2. Game #2: Character Bingo

This game offers a creative variation on a standard bingo game that engages players and teaches the recognition of sounds. This game helps students establish the connection between written characters and auditory sounds, an essential ability for dyslexic children to improve their reading fluency.

Players in this bingo variant hear system-generated audio snippets of spoken characters. Every participant has a bingo card with different Chinese characters on it. They should find and mark the character on their card when they hear its sound. Crossing out a row, column, or diagonal on the card, the player wins the game.

Character Bingo associates the different sounds with the visuals of the characters, amplifying skills in auditory processing while strengthening phonemic awareness. A multisensory way of learning especially helps dyslexic learners in cases where remembering sound-letter correspondence is a struggle. Players get actively involved with learning and facilitate social interactions among peers.

3.5.3. Game #3: Flashcard Frenzy

Flashcard Frenzy is an interactive game in semantic understanding, associated with images of Chinese characters. Games of this nature will help dyslexic students enhance their vocabulary acquisition through visual aids, which is helpful for memory retention and comprehension skills.

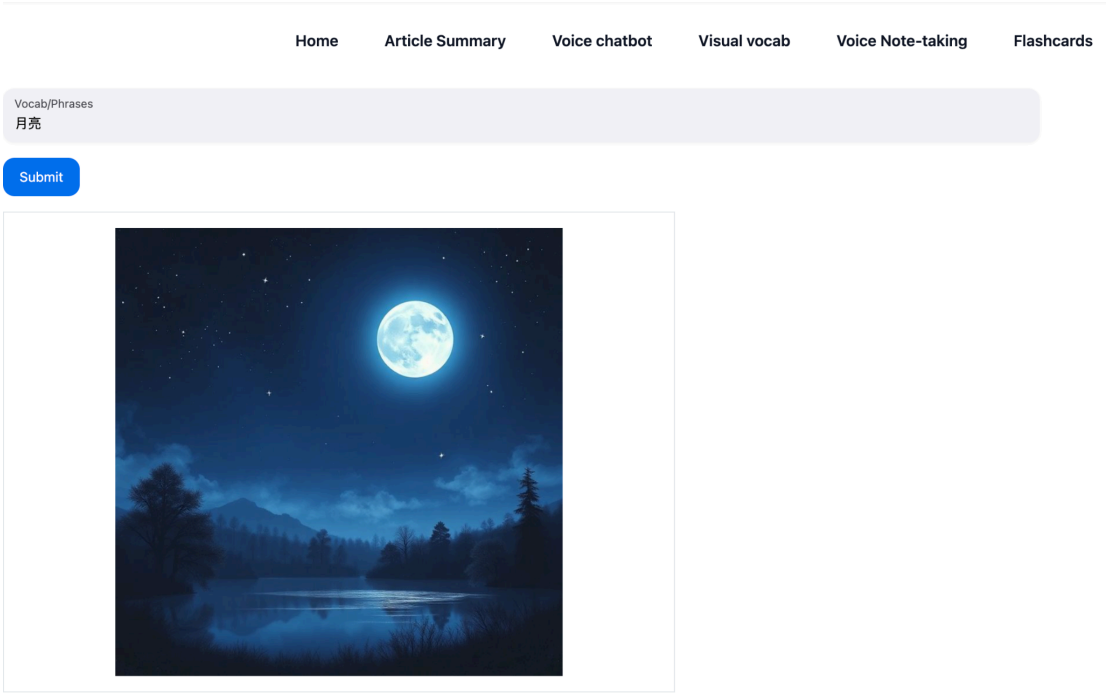
The cards are given in pairs, with pictures on one side and a Chinese character on the other. A player has to find icons that match a given character in the limited time provided. Whenever

the player makes matches, he or she is given instant feedback; this reinforces the correct associations, giving hints for wrong matches.

Flashcard Frenzy lets the learner create meaningful associations between the characters and their meanings, matching the visual stimuli with the text. For dyslexic students, who often have difficulties with conventional rote memorization, this would be particularly effective. Such a game would build up their vocabularies while invoking critical thinking by making them consider word-image relationships. This is because, with repeated play, their recall of characters in context will gradually improve, improving their reading comprehension and expressive language.

4. Current Status

4.1. Vocabulary-to-image tab:

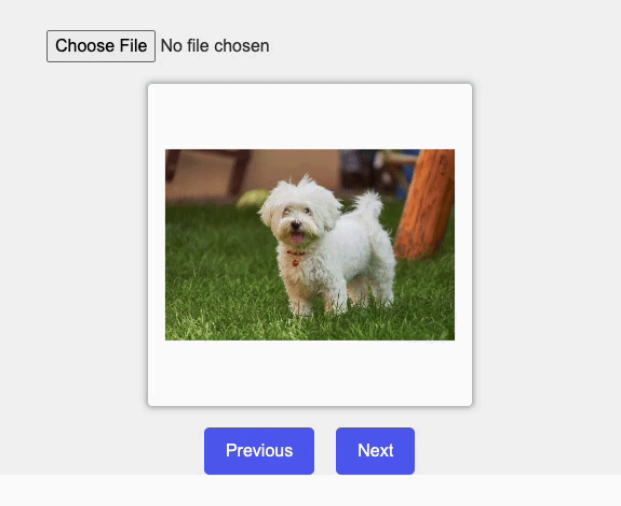


Regarding the vocabulary-to-image function, we have completed the frontend and backend of this feature. The users could input their chosen vocabulary in the input field, a Chinese word for moon in the figure. After clicking the “Submit” button, an image of the moon will be generated to display to the students via communicating with the pollinations.AI in the backend server. A loading sign will be shown if the image is being fetched from the pollination API.

4.2. Voice Chatbot

For the current stage of the voice chatbot, the chatbot could record the student's voice (audio record button) to convert the Cantonese language to written text. The written text could then be sent to the Azure OpenAI for communication, and the response will be returned in the left column. The response text will be automatically read out in Cantonese language using the Azure speech-to-text API.

4.3. Flashcard Generator



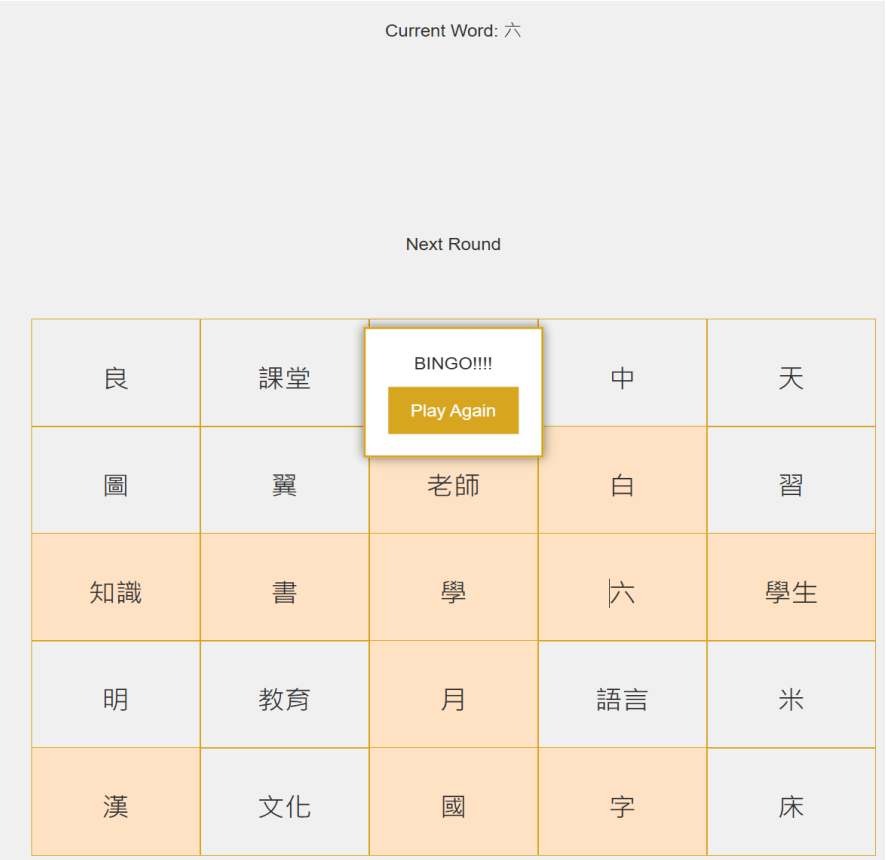
In the current stage of the flashcard generator, we have developed the basic flashcard UI with images as questions and vocabulary as answers. This helps the dyslexic student construct a visual understanding of Chinese characters. The file upload button is expected to accept a short Chinese article and generate flashcards of keywords based on the article, which is done in the backend field of this feature.

Regarding the flashcard generator's backend, we have implemented a backend Python function that automates the extraction of text from images and generates corresponding images for identified keywords. The next steps involve integrating this backend functionality with the frontend user interface, allowing users to upload images and seamlessly receive their flashcards.

4.4. Game #1 (Character Puzzle):

For now, the player wins the game by clicking the correct radical in the right order and then moves on to the next level by clicking the next character.

4.5. Game #2 (Character Bingo):



For now, the character on the bingo card is randomly generated, and the target character/pronunciation will be shown in the *Current Word* field. By getting a row, column, or diagonal in the bingo card, a “Bingo!!!!” will be shown as a winning message.

5. Future Plans

To ensure the ideal, appealing, and successful user learning experience, we have set up a thorough improvement plan to make the frontend and backend features of the all-in-one dyslexia assistance more feasible.

5.1. Major Item

5.1.1. Frontend

5.1.1.1. Refining UI Style

We will work on enhancing the general user interface to make it more visually appealing and easier to use. This includes updating color schemes, typography, and layout designs to improve accessibility and engagement.

5.1.1.2. Iterations of Games

In the current status, the character puzzle game and game puzzle for character introduction or reinforcing it with character bingo or flashcard frenzy activities—the series of gameplay refinement, improvement on art asset display, and response/user action will be realized. This activity will consider valuable contributions from user reviews to better facilitate those games to satisfy their goals, related specifically to the curriculum.

5.1.1.3. Vocabulary revision section

Build a frontend page for students to revise the generated flashcards.

5.1.2. Backend

5.1.2.1. Major planned backend function

- Voice Chatbot: Add a prompt engineering feature in the voice chatbot backend to better deliver human-readable messages to dyslexic students.
- Flash Card Generator: integrate the flashcard with a database to store and get the flashcard details. The user could revise the flashcard generated before.
- Gamified Exercise: Add text-to-speech and text-to-image functions to the character puzzle and character bingo game to help the dyslexic student learn about the visual and auditory connection with text.
- Vocabulary Revision Corner: When the student inputs vocabulary in the vocabulary-to-image tab, a flashcard will be generated to store in the database for future revision.
- Login Page: By including a secure login page, users could register and monitor their progress over time. It improves data management and adjustments.

5.1.2.2. Vector database integration with chatbot

To further enhance the chatbot functionality for the students, we are planning to integrate the chatbot with a vector database as a specific knowledge domain base to provide a more accurate response to the dyslexic student in terms of the learning materials. Students should upload a PDF document (i.e., learning materials) to the chatbot vector database. By breaking down the pdf text into number-like forms (vectors), the chatbot could do a semantic search to find the most correlated information in the database concerning the user query. Therefore, the chatbot could answer the questions according to the learning materials uploaded. The dyslexic student could communicate with the chatbot to understand the learning materials instead of reading all the text of the learning materials.

For the technology tools, we plan to utilize the Azure vector database feature in Azure Cosmos DB to implement the vector database of the chatbot. For the semantic search functionality, we are planning to integrate it with the vector search functionality in the Azure AI search module. Using the above tools, the chatbot could generate an accurate response to the learning materials.

5.2. Future Timeline

January	● Chatbot and flashcard generator frontend-backend integration ● chatbot prompt engineering
February	● Game #1 (Character Puzzle) enhancement ● Login system ● Database design and integration for features
March	● Vector Database for voice chatbot ● Game#2 (Character Bingo) enhancement
April	● Game#3 (Flashcard Frenzy) enhancement ● Final Testing of Dyslexic Assistant System

6. Conclusion

The all-in-one Dyslexia Assistant System offers a comprehensive solution to the fundamental difficulties dyslexic students, in particular, encounter when learning Chinese characters. Based on four main features—auditory-guided visualization, simplifying complex phrases, voice communication, and game-based learning—the development process offers a comprehensive interactive learning environment that can accommodate the various needs of every learner.

Even if the system has advanced significantly, more improvements are anticipated to guarantee its efficacy. These include finishing the creation of the Flashcard Frenzy game, enhancing backend integration for accurate routing, upgrading frontend design, and adding sophisticated capabilities like vector database support for flashcard-to-image generation. These upgrades will increase performance, usability, and engagement, giving students a smooth learning experience.

In addition to enhancing literacy abilities, this approach gives dyslexic children more self-assurance and enthusiasm. The dyslexia helper system is a useful tool that makes learning more inclusive and accessible for everyone by fusing cutting-edge technology with instructional techniques.

7. Reference

Education Bureau. (2021). Types of Special Educational Needs. https://sense.edb.gov.hk/en/types-of-special-educational-needs/index.html

Education Bureau. (2024a). Key Statistics on Primary Education [Excel file]. Retrieved from https://www.edb.gov.hk/attachment/en/about-edb/publication-stat/figures/pri.xlsx

Education Bureau. (2024b) Legislative Council Panel on Education: Professional Support Provided to Ordinary Schools for Implementing Integrated Education.

International Dyslexia Association. (2018). Definition of dyslexia. https://dyslexiaida.org/definition-of-dyslexia