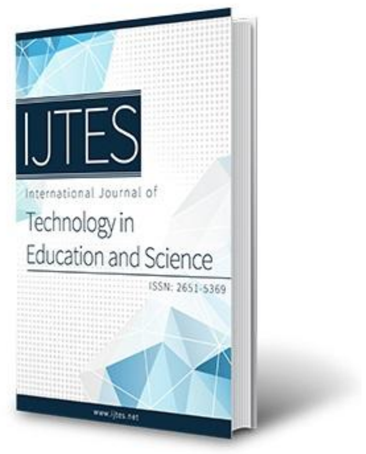

www.ijtes.net

ReadSmart: Generative AI and Augmented Reality Solution for Supporting Students with Dyslexia Learning Disabilities

Yushan Kuerban
University of Exeter, United Kingdom

Solomon Sunday Oyelere
University of Exeter, United Kingdom & University of Johannesburg, South Africa

Ismaila Temitayo Sanusi
University of Eastern Finland, Finland

To cite this article:

Kuerban, Y., Oyelere, S.S., & Sanusi, I.T. (2025). ReadSmart: Generative AI and augmented reality solution for supporting students with dyslexia learning disabilities. *International Journal of Technology in Education and Science (IJTES), 9*(1), 159-176. https://doi.org/10.46328/ijtes.599

International Journal of Technology in Education and Science

2025, Vol. 9, No. 1, 159-176 https://doi.org/10.46328/ijtes.599

ReadSmart: Generative AI and Augmented Reality Solution for Supporting Students with Dyslexia Learning Disabilities

Solomon Sunday Oyelere, Yushan Kuerban, Ismaila Temitayo Sanusi

Article Info

Article History

Received:
29 July 2024
Accepted:
24 December 2024

Keywords

Dyslexia
Learning disabilities
Augmented reality
Generative AI
Educational technology

Abstract

Dyslexia is a learning disability that significantly hinders students' abilities to read and comprehend educational materials, posing a substantial challenge within educational environments. This paper introduces an innovative educational system, ReadSmart, that integrates both Augmented Reality (AR) and Generative Artificial Intelligence (GenAI) technologies, specifically designed to support students with dyslexia. ReadSmart utilizes Microsoft HoloLens 2, Unity3D, and the Mixed Reality Toolkit (MRTK2) to develop an immersive AR environment. Key functionalities of ReadSmart include Optical Character Recognition (OCR) facilitated by Azure Cognitive Services, which extracts text from images captured in real-time, and OpenAI's GPT-4 model, which is responsible for generating concise text summaries and illustrative images to enhance comprehension. Moreover, ReadSmart incorporates Text-to-Speech (TTS) capabilities to improve the accessibility of content for dyslexic learners. Evaluations conducted through self-assessment and peer feedback suggest that ReadSmart has significant potential in reducing the time required for reading and in enhancing comprehension. Initial findings indicate that this system could substantially improve the educational experience for students with dyslexia. Future research will focus on refining ReadSmart's features and conducting comprehensive trials to validate its effectiveness and explore its broader applicability.

Introduction

Dyslexia, a reading disability that is prevalent worldwide, affects a significant proportion of the global population. Research indicates that between 10% and 20% of individuals across different regions experience some form of dyslexia. In the United Kingdom alone, it is estimated that approximately 10% of the population, equating to around 6.7 million people, are affected by dyslexia. Of these, about 20% are school-aged children, and a considerable number remain undiagnosed (Association, 2020). Despite the high prevalence of dyslexia, many individuals remain undiagnosed, and as a result, they do not receive the necessary interventions that could significantly improve their educational outcomes.

Dyslexia is characterized by a spectrum of symptoms that impair an individual's ability to read, write, and process information effectively (Lyon, Shaywitz, & Shaywitz, 2003). Common symptoms of dyslexia include unpredictable and inconsistent spelling, confusion of letter order within words, slow reading speed, frequent errors when reading aloud, difficulty following sequences of instructions, slow writing speed, poor handwriting, and poor short-term memory (NHS, 2022). In addition to these, individuals with dyslexia may also struggle with phonological processing, working memory, and visual processing, which exacerbates their reading difficulties (Ziegler & Goswami, 2005). These challenges can severely impact academic performance, leading to frustration, decreased motivation, and lower self-esteem among dyslexic students (Burden, 2008).

Traditional educational methods and technologies often fall short in addressing the specific and diverse needs of dyslexic learners. Conventional approaches such as specialized tutoring and phonics-based interventions, while helpful, may not offer the comprehensive support required to address the full range of challenges faced by students with dyslexia (Pino & Mortari, 2014). Additionally, these methods often lack the level of personalization and real-time, context-sensitive assistance that could significantly enhance learning outcomes for dyslexic students (Joshi, Dahlgren, & Boulware-Gooden, 2002).

To address these gaps, this project proposes the integration of Augmented Reality (AR) and Generative AI (GenAI) technologies to create a supportive and tailored learning environment specifically for adolescent students with dyslexia. Throughout the content of this study, references to students with dyslexia should be understood to include adolescent students. By leveraging AR to overlay digital content onto real-world environments and utilizing GenAI to generate personalized educational materials, the proposed system aims to enhance the learning experience for dyslexic students. This innovative approach seeks to overcome the limitations of existing educational methods by incorporating advanced technology to create a more effective, engaging, and inclusive educational solution.

Thus, this study is focused on addressing the following research questions:

1. How can AR and GenAI be integrated to design and implement an educational system that effectively supports students with dyslexia?
2. How can user perceptions of usability and functionality inform necessary improvements to the system?

Background

Global Impact of Dyslexia

Dyslexia is a pervasive learning disability that affects millions of people across the globe. In the United Kingdom, approximately 10% of the population is affected by dyslexia, which equates to around 6.7 million individuals (Association, 2020). However, dyslexia is not confined to the UK; it is a widespread issue affecting various regions globally. For instance, in the European Union, the prevalence rate of dyslexia is estimated to be around 12%, reflecting the significant number of students and adults who struggle with reading difficulties (Union, 2021). In Canada, up to 20% of the population is reported to have dyslexia, underscoring the extensive reach of this learning disability (Canada, 2022). Similarly, Australia reports a dyslexia prevalence rate of approximately 10% (Australia, 2023). These statistics highlight the global nature of dyslexia and emphasize the urgent need for effective interventions and support systems to assist those affected by it.

Diagnosing dyslexia remains a complex and challenging process due to the variability in symptoms and the absence of standardized diagnostic criteria (Snowling, 2014). Dyslexia can manifest in various ways, such as difficulties with phonological processing, working memory deficits, and visual processing issues, making it difficult to establish a universal diagnostic approach (Lyon, Shaywitz, & Shaywitz, 2003). Additionally, cultural and linguistic differences can further complicate the diagnostic process. For example, the symptoms of dyslexia may present differently in languages with consistent orthographies compared to those with irregular spelling rules. In many regions, there is a significant lack of awareness and resources dedicated to understanding and diagnosing dyslexia, leading to delayed or missed diagnoses and leaving many individuals without the necessary support to succeed academically and socially (Ziegler & Goswami, 2005).

Impact of Dyslexia on Daily Life and Education

Dyslexia significantly impacts various aspects of an individual's life, particularly in academic settings. Students with dyslexia often encounter challenges in reading comprehension, writing skills, and overall academic performance (Lyon, Shaywitz, & Shaywitz, 2003). These difficulties stem from the inherent challenges that dyslexia poses in processing written language. For instance, dyslexic students may struggle to decode words, leading to slow and laborious reading (NHS, 2022). This difficulty in reading can affect their ability to keep up with reading assignments and to comprehend complex texts, ultimately impacting their performance across all subjects that require reading and writing.

The emotional and social effects of dyslexia are profound. The persistent struggle with reading and writing can lead to feelings of frustration and inadequacy, which contribute to low self-esteem (Burden, 2008). Many dyslexic students experience anxiety, particularly in academic settings where they may be required to read aloud or write in front of peers. This anxiety can extend beyond the classroom, affecting social interactions and relationships (Alexander-Passe, 2006). As a result, dyslexic individuals might avoid social situations where reading or writing is involved, leading to social isolation and difficulties in forming and maintaining friendships (Riddick, Sterling, Farmer, & Morgan, 1999).

The economic and societal impact of dyslexia is considerable. The costs associated with providing educational support for dyslexic students, including specialized tutoring, assistive technologies, and individualized education plans (IEPs), can be substantial (Shaywitz & Shaywitz, 2005). Furthermore, undiagnosed and untreated dyslexia can result in long-term economic consequences. Individuals with dyslexia are more likely to drop out of school, leading to lower employment rates and reduced earning potential (Gerber, 2012). This, in turn, contributes to broader societal costs, including increased reliance on social services and decreased economic productivity (Riddick, 1996).

AR in Education

Augmented Reality (AR) technology overlays digital information onto the physical world, creating an interactive and immersive experience for users (Azuma, 1997). AR systems typically involve the use of devices such as smartphones, tablets, or AR headsets like Microsoft HoloLens, which use cameras and sensors to detect the physical environment and overlay digital content onto it. The software components of AR systems include computer vision algorithms that recognize and track objects in the real world and 3D rendering engines that create and display the digital overlays (Craig, 2013).

AR technology functions by integrating real-time data from the physical environment with computer-generated elements, allowing users to see and interact with digital content in a way that is seamlessly integrated with their real-world surroundings (Azuma, 1997). Various techniques are employed in AR, such as marker-based AR, which uses visual markers like QR codes to trigger the display of digital content, and markerless AR, which relies on the device's sensors to understand the environment and place digital objects within it (Zhou, Duh, & Billinghurst, 2008).

The application of AR in education has demonstrated significant potential in enhancing the learning experience. For example, interactive textbooks that employ AR can bring static images to life, allowing students to scan images in their textbooks and view 3D models, animations, or videos that provide additional context and explanations (Akçayır & Akçayır, 2017; Chang & Hwang, 2018). This interactive element engages students and makes learning more dynamic.By integrating AR technology into the educational process, educators can create more inclusive and effective learning environments that cater to the diverse needs of their students. This innovative approach has the potential to transform traditional educational methods and significantly improve learning outcomes.

GenAI in Personalized Learning

Generative AI (GenAI) refers to a subset of artificial intelligence that uses machine learning models to generate new content based on the data on which it has been trained. These models, such as Generative Adversarial Networks (GANs) and transformer-based models like GPT-4, can produce text, images, and other media types that mimic human-created content. The primary working principle of GenAI involves training on large datasets, learning patterns and structures within the data, and then using this learned knowledge to generate new, coherent outputs that align with input prompts.

In the field of education, GenAI has shown considerable potential in enhancing learning efficiency, particularly for students with reading difficulties such as dyslexia (Holmes, Bialik, & Fadel, 2019). Rather than creating entirely new content, GenAI can summarize and refine existing educational materials, presenting them in a more accessible format. This includes generating related images and videos that visually represent the summarized content, thereby aiding comprehension (Luckin & Holmes, 2016). By transforming complex text into simpler, more digestible formats and providing visual aids, GenAI helps students, especially those with dyslexia, to better understand and retain information, ultimately improving their overall learning experience (Rello & Bigham, 2017).By incorporating GenAI into the educational process, educators can create more effective and personalized learning environments that cater to the diverse needs of their students. This innovative approach has the potential to transform traditional educational methods, making learning more adaptive, engaging, and accessible.

Related Work

Traditional educational strategies have long been employed to support students with dyslexia. These methods typically involve specialized tutoring, where educators trained in dyslexia support work one-on-one with students to address their specific learning challenges (Lyon, Shaywitz, & Shaywitz, 2003). Phonics-based interventions are also widely used, focusing on teaching the relationships between sounds and their corresponding letters to improve reading skills (Torgesen, 2005). Programs like Orton-Gillingham and Wilson Reading System are well-known examples of structured, multisensory approaches that help dyslexic students develop phonemic awareness, decoding skills, and reading fluency (Joshi, Dahlgren, & Boulware-Gooden, 2002).

While these traditional interventions can be effective, they also have limitations. Specialized tutoring and structured programs often require significant time and financial resources, making them less accessible to all students, particularly those from lower-income families (Macdonald, 2010). Moreover, these methods may not fully address the diverse needs of every dyslexic learner, as they often rely on a one-size-fits-all approach that may not be suitable for all students.

In recent years, technological advancements have introduced a variety of tools designed to assist dyslexic learners. For instance, text-to-speech software such as Kurzweil 3000 (Education, 2021) and Read&Write (Read&Write, 2021) allows students to hear written text read aloud, facilitating better comprehension and reducing the cognitive load associated with reading. Reading apps like Learning Ally (Ally, 2024) and Audible (Audible, 2024) provide access to audiobooks, enabling students to engage with literature and educational content through listening rather than reading. Educational games and apps such as Dyslexia Quest (Nessy, 2023) and Nessy Learning (Nessy, 2024) offer interactive and engaging ways for students to practice reading and spelling skills. These tools often incorporate gamification elements to maintain student interest and motivation. Furthermore, they provide immediate feedback, allowing students to learn from their mistakes in real-time.

Recent studies have explored the potential of integrating AR and GenAI to create more effective educational tools. By combining the immersive and interactive capabilities of AR with the adaptive and personalized features of GenAI, it is possible to develop systems that provide real-time, context-sensitive feedback and customized learning experiences (Billinghurst & Duenser, 2012). For example, an integrated AR-GenAI system could capture real-world text through AR and use GenAI to generate summaries, alternative explanations, and visual aids. This approach not only enhances comprehension but also makes learning more engaging for students with dyslexia. Despite the promising potential, research in this area is still in its early stages, and further studies are needed to explore the best practices for effectively integrating these technologies (Pelletier et al., 2022).Although significant advancements have been made in the application of AR and GenAI technologies in education, existing solutions often fail to fully address the unique needs of dyslexic learners. To overcome these limitations, our project integrates AR and GenAI in a novel way, providing personalized and interactive support tailored specifically for students with dyslexia. The following sections detail the design and implementation of our system, highlighting the innovative aspects and technical challenges addressed.

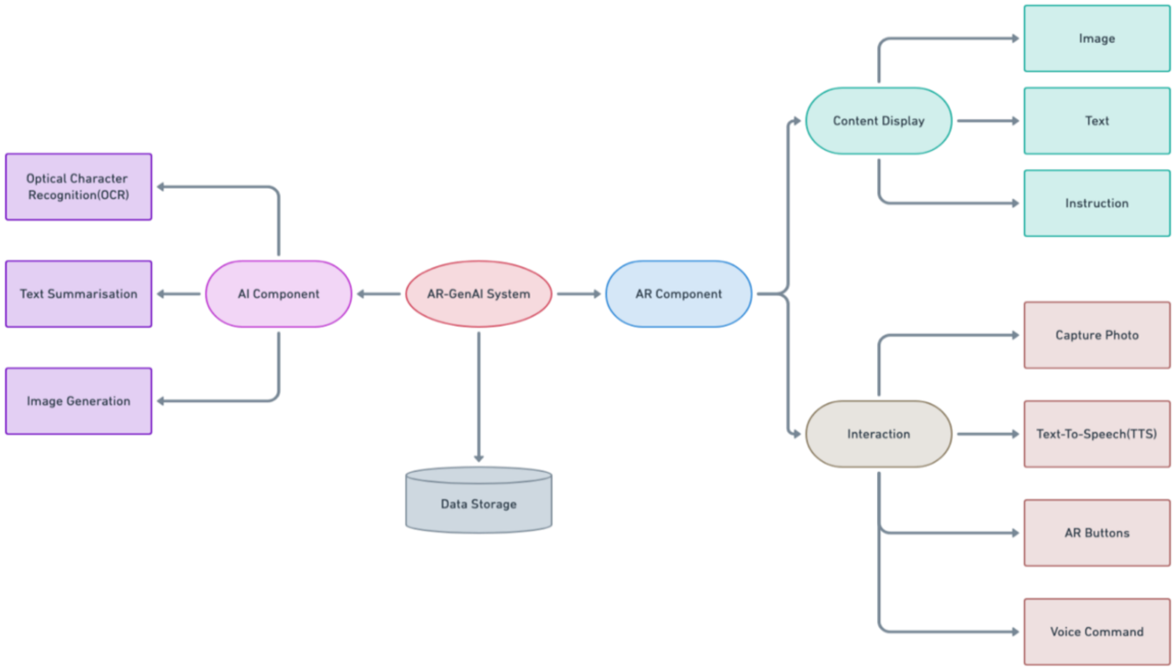

Figure 1. System Architecture

Design and Implementation

System Overview

The primary objective of ReadSmart is to support students with dyslexia by providing an integrated educational platform that leverages both Augmented Reality (AR) and Generative AI (GenAI) technologies. ReadSmart, is designed to enhance the learning experience by making it more interactive and personalized, addressing the unique challenges faced by dyslexic learners. The core functionalities of ReadSmart include capturing text from real-world environments using AR devices, processing the captured text using Optical Character Recognition (OCR), summarizing and generating educational content with GenAI, and displaying this content in an interactive AR environment. Additionally, ReadSmart allows users to interact with it through gestures and voice commands, ensuring a highly engaging and tailored learning experience.

ReadSmart's architecture is designed to facilitate seamless interaction between its various components, ensuring efficient data flow and processing. The architecture comprises the following key modules:

1. Text Capture Module: This module utilizes AR devices to capture text from real-world environments. The captured images are then sent to the Text Recognition Module for processing.
2. Text Recognition Module: This module uses OCR technology to convert the captured images into editable text. The OCR service, such as Azure OCR, can recognize both printed and handwritten text, ensuring high accuracy and versatility.
3. Content Generation Module: Leveraging the capabilities of Generative AI models like GPT-4, this module generates personalized educational content based on the recognized text. The content generated

includes text summaries, explanations, and visual aids tailored to the needs of dyslexic students.

4. Content Display Module: This module presents the generated content within an AR environment. By overlaying digital information onto the real world, it enhances the learning experience through interactive and immersive visuals.

5. User Interaction Module: Facilitating interaction through gestures and voice commands, this module ensures that users can navigate and engage with ReadSmart intuitively. The immediate feedback provided by this module enhances the usability and accessibility of the platform.

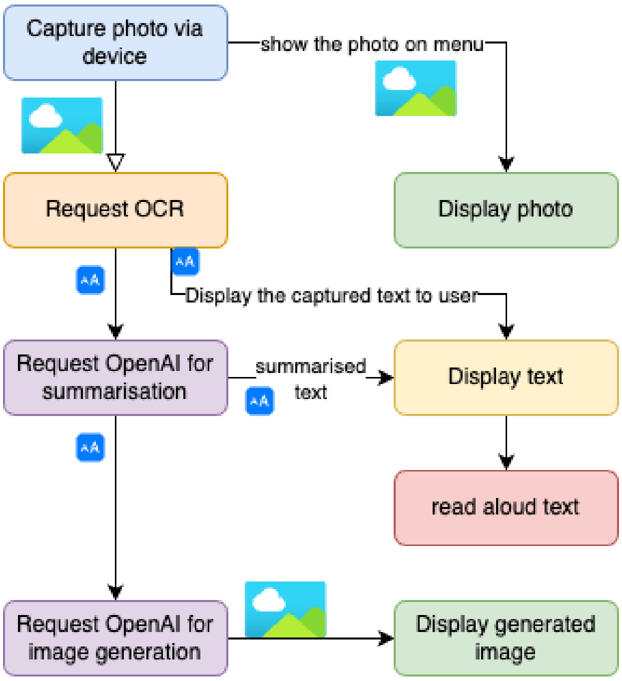

Figure 2. ReadSmart's Interaction Workflow

ReadSmart interaction flow (Fig. 2) illustrates how users interact with ReadSmart. Users can capture text from the real world using the AR device, which is then processed by the OCR service to extract the text. The extracted text is sent to the GenAI model for content generation, which produces personalized educational materials. ReadSmart then displays the generated content in the AR environment, allowing users to interact with it through gestures and voice commands. This interactive and personalized learning experience aims to support dyslexic students in overcoming their reading challenges and enhancing their academic performance.

Technology Stack

The selection of appropriate technologies and tools is critical to the performance and usability of ReadSmart. The chosen technologies include:

- Hardware: The Microsoft HoloLens 2 is selected for its advanced capabilities in capturing and processing real-world images with high precision. Its robust computational power and integration with AR applications make it ideal for this project.

- Software: Unity3D, combined with the Mixed Reality Toolkit (MRTK2), is used for developing the AR environment. Unity3D's versatility and MRTK2's specialized features for AR development ensure a seamless and interactive user experience.
- Programming Language: C is used for the development of ReadSmart due to its strong compatibility with Unity and Azure services. It facilitates the implementation of various functionalities, including user interface design, interaction handling, and integration with cloud services.
- OCR Technology: Azure OCR services are employed for text recognition due to their high accuracy and ability to process various text formats, including handwritten text. This ensures reliable text extraction from diverse sources.
- Generative AI Model: OpenAI's GPT-4 model is utilized for generating personalized educational content. GPT-4's advanced natural language processing capabilities enable it to produce high-quality, contextually relevant educational materials tailored to the needs of dyslexic students.

By integrating these technologies, ReadSmart is designed to provide an effective, engaging, and personalized learning tool for students with dyslexia. The combination of AR and GenAI technologies allows for a seamless blend of real-world interaction and digital enhancement, offering a novel approach to educational support.

User Interaction Design

The design of the user interaction within ReadSmart focuses on creating an intuitive and engaging experience for dyslexic students. Central to this design is the hand menu (handmenu) that follows the user's palm movements, providing an accessible interface for interacting with ReadSmart's functionalities.

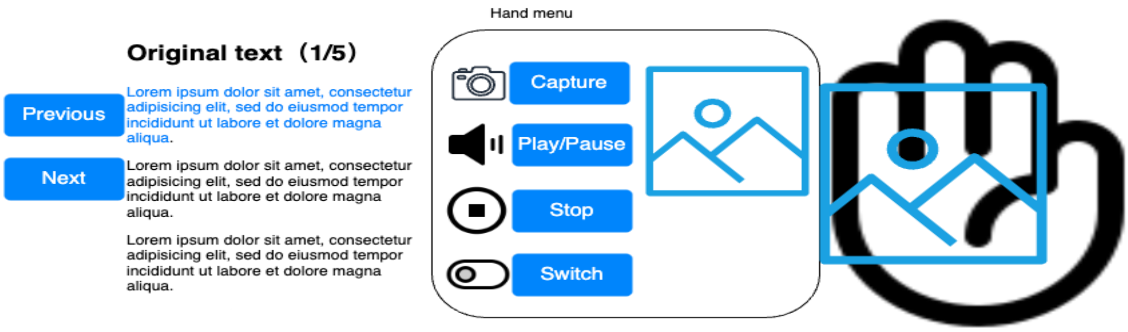

Figure 3. Handmenu Interface Design

The handmenu (Fig. 3) is designed to appear on the inner side of the user's palm when it is within the field of view. As the user moves their palm, the handmenu follows, ensuring that it remains easily accessible. When the user's palm moves out of the field of view, the handmenu stays fixed in its last position using spatial computing techniques, effectively anchoring it in the real world. Additionally, users can drag the handmenu to place it at any desired location within their environment, providing flexibility in interaction.

The handmenu contains four main buttons, each providing a specific functionality, as well as several display areas for showing text, images, and captured content. The key features of the handmenu include:

- Text Display Area: On the left side of the handmenu, there is a text display area that shows either the recognized text or the summarized text, depending on the user's selection. This area also includes two additional buttons for navigating through the history of captured texts. Users can switch to the previous or next text, with ReadSmart supporting a history of up to five recent captures. This feature allows users to review and compare different texts and their summaries. During TTS playback, the text is highlighted in sync with the audio narration, providing a synchronized visual aid.
- Image Display Area: The right side of the handmenu is dedicated to displaying the images generated by ReadSmart based on the summarized text. This visual representation aids in the comprehension of complex concepts, providing a multi-sensory learning experience. The images are updated in sync with the text navigation, ensuring that the visual aids correspond to the selected text.
- Captured Image Preview: A specific area within the handmenu shows a preview of the captured image. This allows users to confirm that they have successfully captured the desired text area before proceeding with text recognition and processing.
- Capture Button: Allows users to capture text from real-world environments using the HoloLens 2 cameras. The captured text is then processed by ReadSmart for further analysis and content generation.
- Play/Pause Audio Button: Enables users to play or pause the audio narration of the recognized text or generated summaries. This feature leverages ReadSmart's TTS capability to provide auditory support.
- Stop Audio Button: Allows users to stop the audio playback, giving them control over the listening experience.
- Toggle Original/Summary Text Button: Enables users to switch between the original recognized text and the summarized version generated by ReadSmart, providing flexibility in content presentation.
- History Navigation Buttons: Enable users to switch to the previous or next text, supporting a history of up to five recent captures. This feature allows users to review and compare different texts and their summaries.

The interaction mechanics are designed to be user-friendly and responsive to ensure that students can focus on learning rather than navigating the interface. The handmenu's ability to follow the user's palm and remain within easy reach reduces the need for extensive hand movements, minimizing physical strain. The spatial anchoring of the handmenu when the palm moves out of view ensures that users do not lose track of the interface, maintaining a seamless interaction flow.

By combining visual, auditory, and tactile interaction methods, ReadSmart accommodates various learning styles and preferences. This multimodal approach ensures that dyslexic students can engage with the educational content in a way that best suits their individual needs, ultimately enhancing their learning experience.

Evaluation and Testing

The evaluation and testing phase is critical to ensure that the AR and GenAI system effectively supports dyslexic students and meets the project's objectives. Given the current constraints, testing primarily involves self-testing and feedback from peers, focusing on functionality, performance, usability, and ReadSmart's ability to save reading time.

The performance testing was conducted in two main phases. In the first phase, ReadSmart was subjected to 50 tests using the same set of text data. This phase focused on assessing ReadSmart's response time, content accuracy, and the quality of generated images. Key metrics, such as average processing time per task and accuracy rates in text recognition and summarization, were meticulously recorded.

In the second phase, ReadSmart's robustness was tested using a diverse set of 50 text samples, including various fonts and handwritten texts. This phase aimed to evaluate ReadSmart's adaptability and accuracy across different types of input, simulating real-world usage conditions. The testing included measuring the time taken to process each type of text and the accuracy of the output, particularly focusing on how well ReadSmart handled more challenging inputs, such as poorly legible handwriting.

To evaluate ReadSmart's usability and overall user experience, a comprehensive online user testing approach was adopted. Participants, including individuals with backgrounds in education, such as students and professors, as well as those from other industries, were invited to participate in the online survey. A total of 45 professors/teachers and 258 students participated in the survey. The demonstration video provided a first-person perspective on ReadSmart's functionalities, including text capture, summary generation, image creation, and the TTS feature.

After watching the video, participants were asked to complete a standardized survey designed to gather both qualitative and quantitative feedback. The survey included questions aimed at assessing the ease of use, intuitiveness of navigation, and overall user interface design. Participants were prompted to rate their agreement with statements regarding the clarity and user-friendliness of the interface, as well as any points of confusion or difficulty they perceived.

Additionally, the survey collected feedback on the quality and relevance of the generated summaries and images. Participants evaluated how well these outputs aligned with the original content and whether they found the visual aids helpful in understanding the material. The effectiveness of the TTS feature was also reviewed, focusing on the clarity of the synthesized voice, the pace of speech, and its overall utility in enhancing the learning experience. The insights gathered from this online user testing approach are invaluable in identifying ReadSmart's strengths and areas for improvement. These insights are crucial for guiding future iterations of ReadSmart, ensuring that it meets the needs of its intended users and provides a seamless and supportive learning experience. Based on the feedback and data collected from these evaluations, ReadSmart undergoes iterative refinement. Adjustments are made to address identified issues, enhance functionalities, and improve the overall user experience. This iterative approach ensures that ReadSmart evolves based on practical feedback and continues to meet the educational needs of dyslexic students effectively. Although direct testing with dyslexic students was not feasible, the evaluation and testing phase involving detailed peer feedback is essential for validating ReadSmart's technical performance and user experience. This approach ensures that the project delivers a robust, user-friendly, and effective educational tool that can significantly enhance the learning experience for dyslexic students.

Results

System Performance Evaluation

ReadSmart's performance evaluation focuses on assessing the functionality and efficiency of the AR and GenAI system based on extensive testing across various components, including OCR accuracy, text processing and generation speed, overall response time, and system stability.

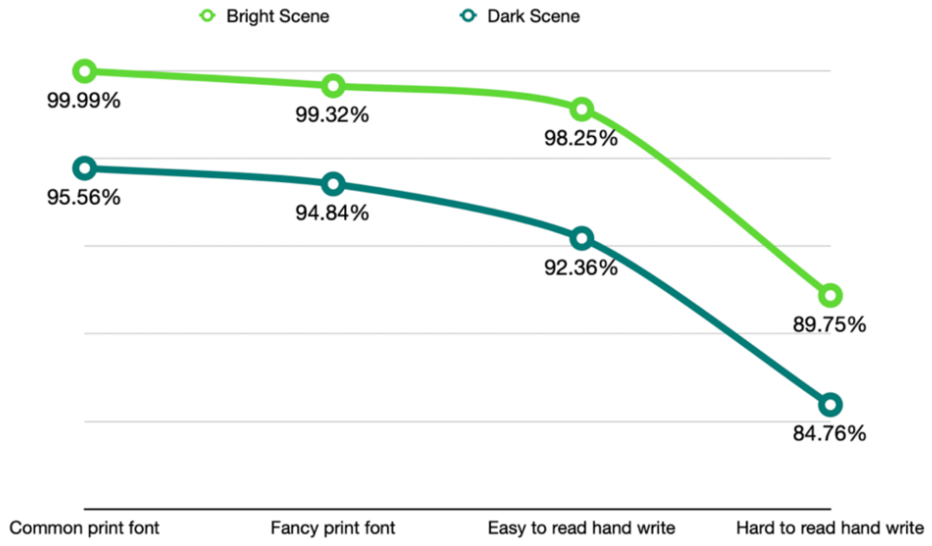

Figure 4. OCR Accuracy Test Result

The OCR component demonstrated high accuracy over a total of 100 tests conducted under different lighting conditions and with various text types. In bright scenes, the OCR accuracy for common print fonts was exceptionally high at 99.99%, representing the best performance of ReadSmart. When processing more ornate print fonts, the accuracy slightly decreased to 99.32%. For easily readable handwritten texts, the accuracy remained robust at 98.25%, but it dropped to 89.75% when handling difficult-to-read handwriting.

As the chart shows, the OCR performance exhibits a slight decline under darker lighting conditions, with the average accuracy dropping by 4-5% for all text types. Despite the lower accuracy rates in certain challenging scenarios, such as difficult handwriting and low-light conditions, these situations had a minimal impact on the overall testing process. This is because ReadSmart is primarily designed to handle common printed text, which constitutes most use cases. Thus, the overall reliability and practicality of ReadSmart remain intact for its intended applications.

In terms of average response time, a cycle of text capture, processing, and content generation took an average of 5.5 seconds, with image generation processing accounting for most of the time. The capture and OCR components were highly efficient, taking only 2 seconds on average, while the GenAI model processing time varied depending on the complexity of the text and the length of the summary generated. About 64% of the total processing time was spent on text and image generation, indicating potential areas for optimization in this component.

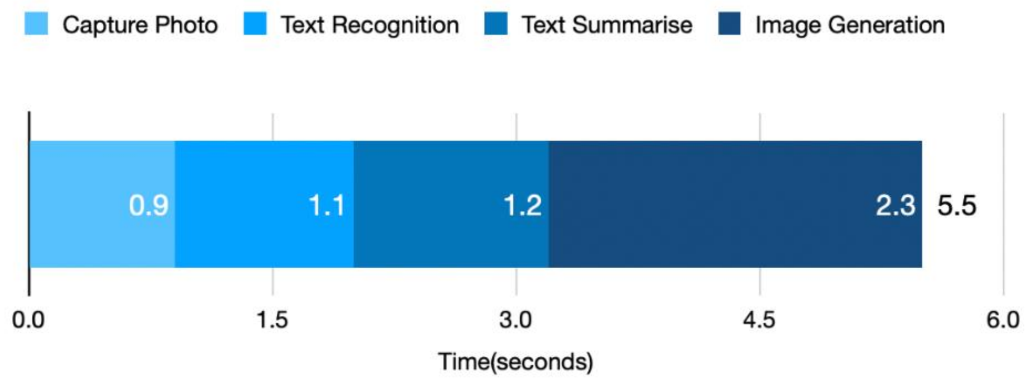

Figure 5. Average Processing Time for Each System Component

Based on self-testing and peer feedback, the reading time was significantly reduced when using ReadSmart compared to traditional reading methods. It took approximately 3-5 minutes to read and understand a 500-word text before using ReadSmart, whereas using ReadSmart reduced the time to 1-2 minutes. This reduction in reading time was attributed to ReadSmart's summarization and image generation features, which provided concise and visually engaging content that facilitated quicker comprehension.

Regarding system stability, the platform demonstrated robust performance throughout the testing phase. Earlier issues related to performance degradation and occasional crashes were addressed through targeted optimizations, resulting in a stable operational environment. The absence of recent performance issues highlights the effectiveness of these optimizations, ensuring reliable system performance under typical usage conditions.

User Feedback

During the testing phase, user feedback was collected from over 40 participants, including 4 professors/teachers, 25 students, and 13 professionals from non-education fields, through an online survey that included a detailed video demonstration of the AR system. The primary feedback revolved around the user interface, interaction experience, and overall system functionality.

The user testing phase involved a detailed survey designed to gather both quantitative and qualitative feedback on ReadSmart's usability, user interface, content generation, and TTS features. Participants provided ratings on various aspects of ReadSmart, and their comments offered deeper insights into their experiences. After collecting the survey responses, the data was analyzed to identify key trends and patterns in the feedback. The weighted mean for each question and the standard deviation were calculated to assess the overall user satisfaction and identify areas of consensus and divergence.

Quantitative data from the survey revealed that ReadSmart received an overall satisfaction rating of 3.9 out of 5 from users, indicating a generally positive reception. The standard deviation for this question was 0.63, suggesting that user feedback was relatively consistent, with most users providing scores close to the mean. This low standard deviation reflects a consensus on ReadSmart's overall effectiveness, with few outliers indicating significant dissatisfaction.

The TTS feature received a high mean score of 4.1, reflecting that most users found this feature beneficial for aiding comprehension. However, the standard deviation for this feature was 0.85, indicating some variability in user experiences. This suggests that while many users rated the TTS feature highly, there were differing opinions, possibly due to individual preferences or variability in how the feature performed across different scenarios.

Table 1. Result of System Usability Survey

Features	Mean	SD
Overall satisfaction	3.9	0.63
UI is intuitive and user-friendly	3.7	0.86
Navigation is straightforward	3.7	0.82
Response time is satisfactory	3.7	0.85
OCR accuracy meets expectations	4.0	0.82
TTS feature helps understanding	4.1	0.85
Visual content is helpful	3.9	0.79
Clear understanding from the demo	4.1	0.67
Enhances reading comprehension	3.7	0.66

SD: Standard Deviation

For OCR accuracy, ReadSmart achieved a mean rating of 4.0. This score suggests that users generally felt ReadSmart met their expectations in accurately recognizing and processing text, which is crucial for dyslexic students who rely on precise text interpretation. The standard deviation here was 0.82, showing that while most users had similar positive experiences with the OCR functionality, there was some level of disagreement, potentially pointing to occasional inconsistencies in text recognition.

The user interface (UI) was rated at 3.7, indicating that while users found the interface somewhat intuitive, there is room for improvement. The standard deviation for the UI rating was 0.86, which is relatively high, suggesting that user opinions varied significantly. Some users may have found the interface easy to use, while others may have encountered difficulties, highlighting a need for further refinement in UI design to achieve a more universally positive user experience.

Navigation within ReadSmart also received a rating of 3.7, with a standard deviation of 0.82. The mean score suggests that users generally found the navigation functional, though not without some challenges. The standard deviation indicates that user experiences with navigation were somewhat inconsistent, which might point to areas where the flow or layout could be optimized to improve overall usability. ReadSmart's response time had a mean rating of 3.7, reflecting that users generally found ReadSmart's performance satisfactory, though not exceptional.

The standard deviation was 0.85, indicating that some users experienced delays that negatively impacted their experience, suggesting that improving response time could enhance user satisfaction.
In terms of visual content, users provided a mean rating of 3.9, suggesting that they found the visual aids helpful in supporting learning. The standard deviation for this feature was 0.79, indicating a moderate level of agreement among users, though some may have found the visual content less effective or relevant, which could indicate an area for potential improvement. Finally, the clear understanding from the demo was rated highly, with a mean score of 4.1, reflecting that the demonstration effectively conveyed ReadSmart's functionalities to users. The standard deviation for this aspect was 0.67, indicating a strong consensus among users that the demo provided a clear and understandable overview of ReadSmart.

Qualitative feedback from the open-ended question of the survey complemented these ratings, with many users expressing satisfaction with the project's goals and the innovative integration of AR and GenAI technologies. One participant noted, "The TTS feature is very clear and significantly aids in understanding the text," while another appreciated the OCR functionality for its accuracy. However, significant criticism was directed at the user interface, with comments such as "The UI is cluttered and complex" and "Navigation needs to be more straightforward." Users also highlighted the need for a more streamlined and aesthetically pleasing design, suggesting improvements in user interaction mechanisms and overall visual appeal.

Conclusion and Future Work

Conclusions

The results of this project confirm that the research questions have been effectively addressed. By integrating AR and GenAI technologies, ReadSmart has demonstrated its potential to create an educational platform that supports dyslexic students in a meaningful way. The evaluation indicates that ReadSmart not only meets its intended goals but also offers valuable insights into usability and functionality, which will inform future improvements. These findings underscore ReadSmart's capacity to contribute to more inclusive and effective learning environments while also highlighting areas for ongoing refinement.

ReadSmart successfully integrated key technologies, including Microsoft HoloLens 2 for augmented reality, Unity3D and MRTK2 for development, Azure Cognitive Services for OCR, and OpenAI's GPT-4 for content generation. The inclusion of a dynamic handmenu provided a unique interface for users to interact with ReadSmart, capturing text, receiving summaries, and accessing visual aids. The TTS functionality further enriched ReadSmart by offering an auditory learning component, which was particularly well-received by users. Through testing and feedback, ReadSmart demonstrated high accuracy in text recognition and efficient processing capabilities. Users appreciated the quality of the generated summaries and images, noting that these features significantly aided their understanding of complex texts. The positive reception of the TTS feature highlighted its importance in providing accessible content, further validating ReadSmart's design choices.

However, the feedback also revealed areas where the user experience could be improved, particularly in the intuitiveness of the interface and the physical ergonomics of using the handmenu. These insights are crucial for guiding future iterations of ReadSmart, emphasizing the need for more user-friendly interaction mechanisms and better onboarding processes. This project has developed an innovative tool that meets the specific needs of dyslexic students by integrating advanced technologies. ReadSmart's ability to provide real-time text processing, summaries, and multimodal content delivery represents a significant advancement in educational technology for individuals with learning disabilities. This work lays a strong foundation for future research and development, offering valuable lessons and clear pathways for enhancement.

Limitations

The AR and GenAI system for supporting dyslexic students, while innovative and promising, has several limitations that were identified during the development and testing phases. These limitations highlight areas for improvement and provide directions for future development. One of the primary limitations lies in the user experience and interface design. Feedback from users indicated that finding and operating the handmenu was initially challenging, particularly due to the lack of intuitive guidance and clarity in on-screen prompts. This suggests that the current interface design may not be sufficiently user-friendly, especially for first-time users. Additionally, the reliance on hand gestures for interacting with ReadSmart led to physical discomfort and fatigue, particularly during extended use. This ergonomic issue points to a need for more diverse interaction methods that reduce physical strain.

Technically, while ReadSmart's OCR functionality generally performed well, it struggled with certain inputs, such as poorly handwritten text or text captured in low-light conditions. These scenarios led to decreased accuracy and reliability, indicating a limitation in handling diverse and less-than-ideal input conditions. Similarly, although the TTS feature received positive feedback, its naturalness was rated as moderate, which may affect user engagement and satisfaction during prolonged use.ReadSmart's dependency on advanced hardware, such as the Microsoft HoloLens 2, also presents a significant limitation. The high cost and accessibility of such devices limit ReadSmart's potential user base, making it less feasible for widespread adoption, particularly in resource-constrained educational environments.

Moreover, the current testing scope was limited to a small group of users who provided feedback based on an online demonstration, primarily consisting of non-target users. The absence of direct interaction with dyslexic students, the primary target audience, limits a comprehensive evaluation of ReadSmart's effectiveness and usability in real-world educational settings. This limitation highlights the need for future testing with a broader and more representative user base, including hands-on trials with dyslexic students.

Additionally, while ReadSmart demonstrates strong capabilities in text processing and content generation, its overall functionality remains relatively constrained. The lack of support for voice commands or remote control limits the flexibility and accessibility of user interactions. Furthermore, ReadSmart does not yet fully leverage machine learning for personalized learning experiences, which could adapt to individual users' needs and learning styles over time. These limitations provide a roadmap for future enhancements and refinements. Addressing these issues will be crucial in improving ReadSmart's usability, accessibility, and overall effectiveness, thereby better serving the educational needs of dyslexic students and potentially broader user groups.

Future Research Directions

The future research directions for the AR and GenAI system are focused on enhancing user experience, expanding functionality, and increasing accessibility across various augmented reality hardware platforms. As AR technology continues to evolve, ReadSmart aims to adapt and improve in several key areas.

Firstly, enhancing user experience and interaction methods remains a priority. The introduction of voice commands and remote-control devices, such as handheld controllers, could significantly reduce the reliance on hand gestures, thereby alleviating physical fatigue associated with prolonged use. This development would offer users more flexible and accessible interaction options, catering to different user preferences and physical capabilities. Additionally, refining the user interface to be more intuitive and user-friendly is crucial. This could involve more comprehensive onboarding instructions, interactive tutorials, or even an AR virtual assistant to guide users through ReadSmart's features, particularly for first-time users.

Secondly, in terms of technical performance, there is a need to optimize the OCR and TTS technologies. While the current system performs well, further improvements in the accuracy of OCR, especially in challenging conditions such as handwritten text or low-light environments, would enhance reliability. Similarly, increasing the naturalness and quality of TTS output would improve the overall user experience, making the auditory aspect of ReadSmart more engaging and easier to understand. Incorporating machine learning algorithms could also enable personalized learning experiences, allowing ReadSmart to adapt content presentation and interaction methods based on user feedback and interaction patterns, thus better meeting the diverse needs of individual learners.

Expanding ReadSmart's functionality and application scope is another critical direction. Supporting a broader range of AR glasses beyond the Microsoft HoloLens 2 would make ReadSmart more accessible as AR technology becomes more widespread and affordable. This would involve ensuring compatibility with different AR hardware, thus expanding the potential user base and applicability of ReadSmart. Additionally, extending ReadSmart's capabilities to support multiple languages, including OCR, TTS, and content generation, would cater to a more diverse, multilingual audience, further broadening ReadSmart's reach and utility.These future research directions are aimed at addressing current limitations, enhancing user satisfaction, and ensuring that ReadSmart remains at the forefront of educational technology for dyslexic students and potentially other user groups. By continually evolving and incorporating cutting-edge advancements, ReadSmart can better serve its users and adapt to the changing landscape of AR and educational technologies.

References

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational research review*, *20*, 1–11.

Alexander-Passe, N. (2006). How dyslexic teenagers cope: an investigation of self-esteem, coping and depression. *Dyslexia, 12*, 256–275.

Ally, R. (2024). Reading Ally: A comprehensive reading solution for students with dyslexia. *Reading Ally: A comprehensive reading solution for students with dyslexia*. Retrieved from https://www.readingally.org/

Association, B. D. (2020). *Dyslexia Factsheet*. Retrieved from Dyslexia Factsheet: https://www.bdadyslexia.org.uk/resources

Audible. (2024). Audible: Listen to books, podcasts, and more. *Audible: Listen to books, podcasts, and more*. Retrieved from https://www.audible.com/

Australia, H. D. (2023). *Learning Disabilities*. Retrieved from Health Direct: https://www.healthdirect.gov.au/learning-disabilities

Azuma, R. T. (1997). A survey of augmented reality. *Presence: teleoperators & virtual environments, 6*, 355–385.

Billinghurst, M., & Duenser, A. (2012). Augmented reality in the classroom. *Computer, 45*, 56–63.

Burden, R. (2008). Is dyslexia necessarily associated with negative feelings of self-worth? A review and implications for future research. *Dyslexia, 14*, 188–196.

Canada, D. (2022). *statistic data of dyslexia in Canada*. Retrieved from statistic data of dyslexia in Canada: https://www.dyslexiacanada.org/en/dyslexia-basics

Chang, S.-C., & Hwang, G.-J. (2018). Impacts of an augmented reality-based flipped learning guiding approach on students' scientific project performance and perceptions. *Computers & Education, 125*, 226–239.

Craig, A. B. (2013). Understanding augmented reality: Concepts and applications.

Education, K. (2021). *Kurzweil 3000 Powerful reading, writing, studying and test-taking tools*. Retrieved from https://www.kurzweiledu.com/

Gerber, P. J. (2012). The impact of learning disabilities on adulthood: A review of the evidenced-based literature for research and practice in adult education. *Journal of Learning Disabilities, 45*, 31–46.

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning.* Center for Curriculum Redesign.

Joshi, R. M., Dahlgren, M., & Boulware-Gooden, R. (2002). Teaching reading in an inner city school through a multisensory teaching approach. *Annals of Dyslexia, 52*, 229–242.

Luckin, R., & Holmes, W. (2016). Intelligence unleashed: An argument for AI in education.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of dyslexia, 53*, 2.

Macdonald, S. J. (2010). Towards a social reality of dyslexia. *British Journal of Learning Disabilities, 38*, 271–279.

Nessy, L. (2023). *Dyslexia Quest: A fun game to help children with dyslexia*. Retrieved from https://www.nessy.com/en-gb/product/dyslexia-quest-home

Nessy, L. (2024). *Nessy Learning: Literacy support for dyslexia that follows the Science of Reading*. Retrieved from https://www.nessy.com/uk/

NHS. (2022). *Symptoms of dyslexia*. Retrieved from Symptoms of dyslexia: https://www.nhs.uk/conditions/dyslexia/

Pelletier, K., Mccormack, M., Reeves, J., Robert, J., Arbino, N., Dickson-Deane, C., . . . others. (2022). *2022 educause horizon report teaching and learning edition.* Tech. rep., EDUC22.

Pino, M., & Mortari, L. (2014). The inclusion of students with dyslexia in higher education: A systematic review using narrative synthesis. *Dyslexia, 20*, 346–369.

Read&Write. (2021). *Read&Write Literacy software for reading, writing, and research*. Retrieved from https://www.texthelp.com/en-gb/products/read-and-write-education/

Rello, L., & Bigham, J. P. (2017). Good background colors for readers: A study of people with and without dyslexia. *Proceedings of the 19th international ACM SIGACCESS conference on computers and accessibility*, (pp. 72–80).

Riddick, B. (1996). *The Social and Emotional Consequences of Specific Learning Difficulties.* Routledge.

Riddick, B., Sterling, C., Farmer, M., & Morgan, S. (1999). Self-esteem and anxiety in the educational histories of adult dyslexic students. *Dyslexia, 5*, 227–248.

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological psychiatry, 57*, 1301–1309.

Snowling, M. J. (2014). Dyslexia: A language learning impairment. *Journal of the British Academy*, *2*, 43–58.

Torgesen, J. K. (2005). Recent discoveries from research on remedial interventions for children with dyslexia. *The science of reading*, 521–537.

Union, E. P. (2021). National Dyslexia Measurements Across Europe Report. *Dyslexia Compass*, 12.

Zhou, F., Duh, H. B.-L., & Billinghurst, M. (2008). Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, (pp. 193–202).

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological bulletin, 131*, 3.

Author Information

Yushan Kuerban
https://orcid.org/0009-0003-2485-6073
Department of Computer Science
University of Exeter
Stocker Rd, Exeter EX4 4PY
United Kingdom

Ismaila Temitayo Sanusi
https://orcid.org/0000-0002-5705-6684
School of Computing
University of Eastern Finland
P.O.Box 111, 80101, Joensuu
Finland

Solomon Sunday Oyelere
https://orcid.org/0000-0001-9895-6796
Department of Computer Science
University of Exeter
Stocker Rd, Exeter EX4 4PY
United Kingdom
&
Research Group on Data, Artificial Intelligence, and Innovations for Digital Transformation
Johannesburg Business School (JBS)
University of Johannesburg, Johannesburg
South Africa
Contact e-mail: *s.oyelere@exeter.ac.uk*