# ../../../../../../../../../../Users/andreacurley/Desktop/Screen%20

# MSc in Computer Science - Team Project

# Interim Report Template

# Team Name:

## Introduction

In today's digital age, the concept of accessibility has extended beyond physical spaces to include digital platforms, making it an imperative need. The Accessibilator project was born out of this necessity, aiming to bridge the gap between standard document formats and the unique needs of individuals with learning disabilities, with an initial focus on Dyslexia. A neurological condition that affects roughly 10% of the global population and presents unique challenges in areas such as phonology, orthography, and syntax. This often results in prolonged reading times and reduced comprehension, highlighting the need for specialised document formatting. The project ideally aligns with the theme of "social good," as it not only enhances individual user experience but also contributes to creating a more inclusive digital society.

The primary function of the Accessibilator is to transform a variety of documents—ranging from PowerPoint presentations to Word documents and Excel spreadsheets—into formats that are highly accessible. Upon uploading a document, the system analyses it for potential issues and proactively suggests actionable improvements. Users have the final say in adopting these suggestions, ensuring that the document aligns better with their reading preferences.

Compliance with existing legal frameworks like the Americans with Disabilities Act (ADA), the European Accessibility Act, and the United Nations Convention on the Rights of Persons with Disabilities is a cornerstone of this project. The application aims to not just meet but understand the implications of these laws, thereby setting a standard in digital accessibility.

This report aims to offer a comprehensive overview of the Accessibilator, detailing its objectives, methodologies, and future directions. It serves as a lens into the world of digital accessibility, providing valuable insights into how technology can be leveraged for social good.

Dyslexia, a neurological disability, affects approximately 10% of the global population. This condition challenges individuals in multiple linguistic areas, such as phonology, orthography, morphology, lexicon, syntax, and discourse. Individuals with dyslexia often experience prolonged reading times and reduced comprehension levels, leading to the need for repeated readings. Despite the prevalence of dyslexia, there remains a lack of comprehensive guidance in document content and formatting that caters to this demographic. This gap highlights the need for technological advancements and improved document formatting strategies to enhance the accessibility of online materials.

The primary goal of this project is to develop a specialized website designed to address the unique needs of individuals with dyslexia, with a central focus on enhancing the readability of their documents. Recognizing the diverse range of documents people interact with daily, the platform will accommodate multiple file types, including docx, ppt,xls . After content is uploaded, the system will carefully analyze it to find any accessibility issues that could impair reader’s ability to comprehend it. The platform will proactively suggest actionable improvements, ensuring that the content aligns better with the reading preferences of those with dyslexia. Users are given the authority to make final decisions after the suggested improvements are presented. They can choose to adopt the adjustments selectively in accordance with their own preferences.

The document will be displayed in an interactive viewer on the platform after optimization. Users can further modify the document's appearance and organization using this dynamic interface to make sure it precisely suits their reading comfort. When users are satisfied with the changes, they can easily download the revised document ensuring they have a version tailored for optimal readability, which they can use across various platforms and devices.

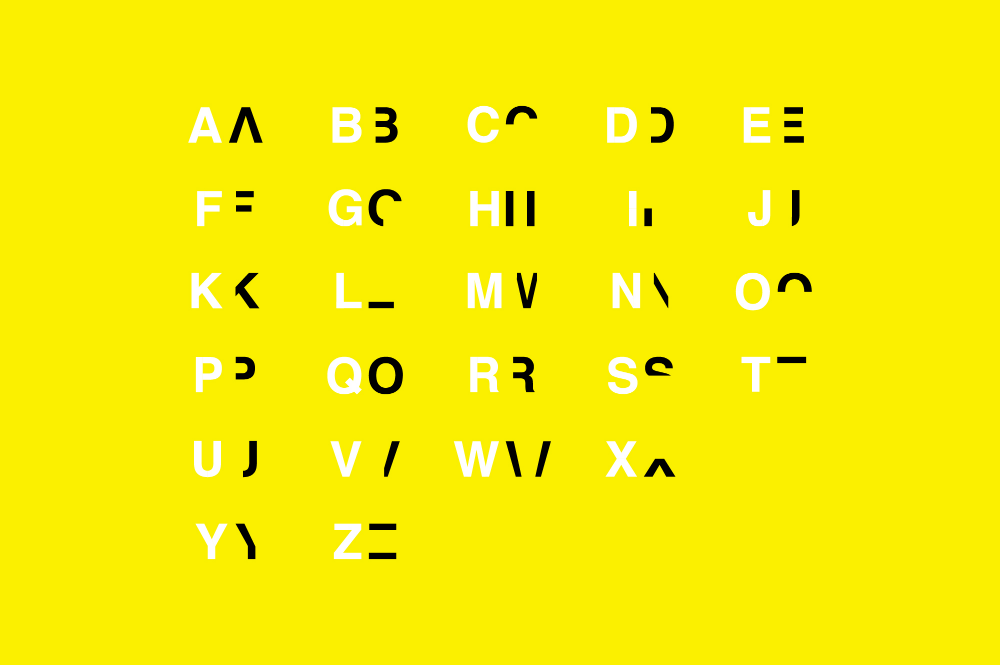
## 1. User Scenario: The Characters (500 words approx.)

* Who is your target user? User personas (diagram,description. How we identified them)
* Why are they important?
* What problem are you solving for them?

The two primary target user groups we identified are people living with dyslexia and people who may know/care for them. These user groups help in acknowledging and comprehending the unique demands of individuals who have dyslexia as well as those who may know or assist them. Our designated user group consists of members from the Dyslexia Association of Ireland. We have maintained ongoing communication with them to ensure their active involvement through surveys and feedback, in the development and testing of our system.

People with dyslexia make up a substantial section of the online user base due to the ubiquitous nature of the disorder. By giving accessibility for users who are dyslexic priority, we help to create a more democratic online environment. It makes sure a larger audience can access information and services, promoting inclusivity and equitable opportunity for everyone.

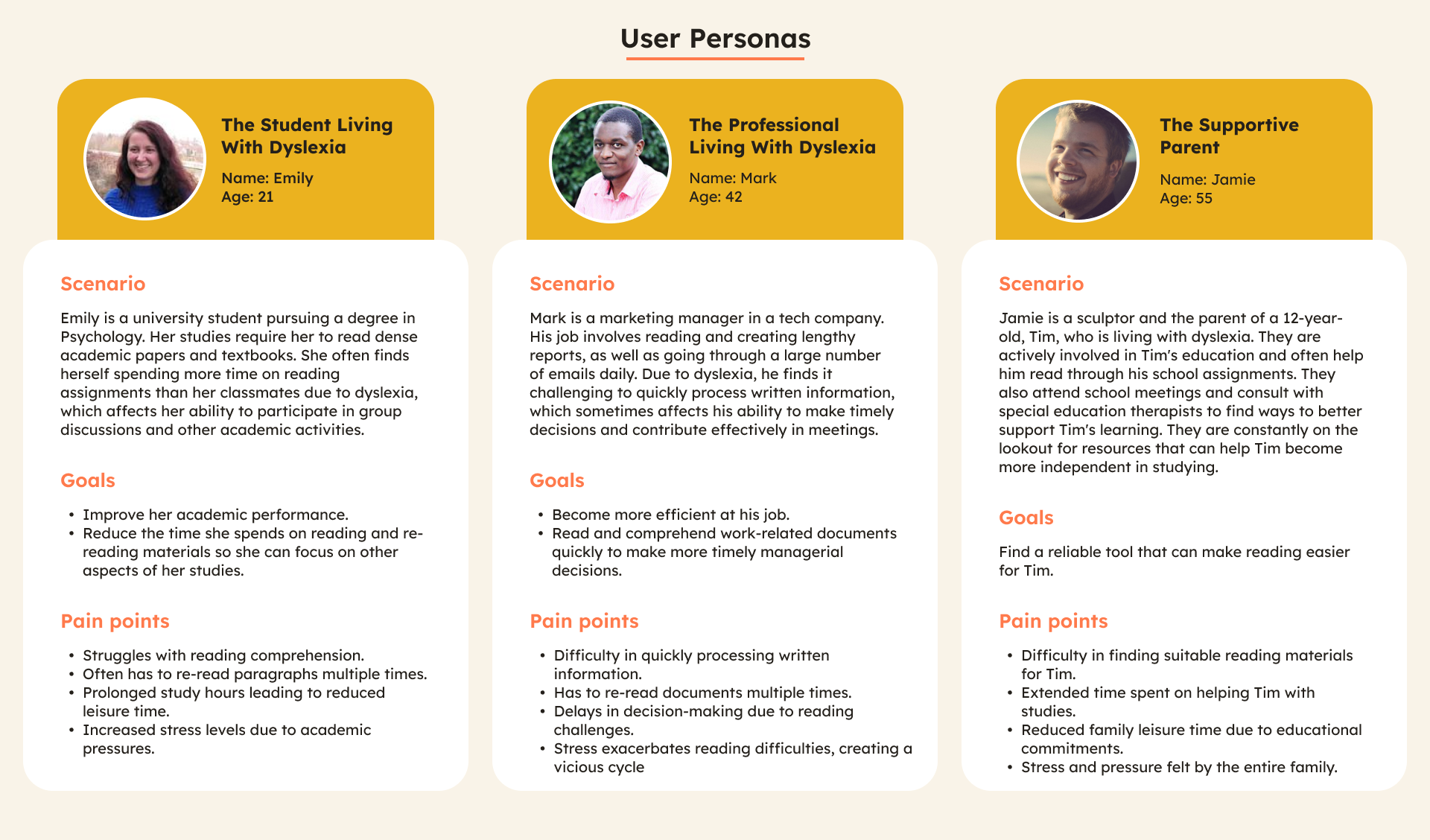
Participants in a study by Grusky, Taft, Naaman, and Azenkot at Cornell Tech, particularly those who had dyslexia, used practices of skipping and skimming sections of text that posed comprehension challenges, as well as the act of re-reading to solidify understanding. The survey participants with dyslexia consistently reported higher frequencies of these behaviours in comparison to their counterparts without dyslexia.[5]



[2] <https://danielbritton.info/dyslexia/>

The consequences of dyslexia extend beyond academic difficulties. Individuals with dyslexia often experience prolonged reading times and reduced comprehension levels, leading to the need for repeated readings. This struggle can lead to heightened levels of stress, anxiety, and depression. Moreover, the societal stigma surrounding dyslexia can exacerbate emotional and psychological burdens, affecting self-esteem and motivation We want to significantly improve this population's reading experience using various document optimization techniques

Our website provides a user-friendly platform for individuals to upload documents. Our system thoroughly analyzes the document after uploading it to find any accessibility issues that may be there. The relevant adjustments are then automatically made for optimal readability.



**User Scenario: The Characters**

**Who is Your Target User? User Personas**

Our target users primarily fall into two categories: individuals living with dyslexia and those who support them. We have identified these user groups through ongoing communication with members of the Dyslexia Association of Ireland and through a review of existing literature on dyslexia and digital accessibility

(Your personas would be inserted here, as you've already finalized them.)

**Why Are They Important?**

Understanding the needs of these specific personas is crucial for the success of the Accessibilator. Emily and Mark represent a significant portion of the dyslexic population who are striving for academic and professional success. Their challenges are not just personal but also represent a broader issue of accessibility and inclusivity in educational and professional environments. By addressing their needs, we are taking a step towards a more equitable society where information is accessible to all, regardless of their neurological conditions

Jamie, on the other hand, represents the caregivers and supporters of those with dyslexia. By making resources more accessible for them, we indirectly improve the quality of life for the dyslexic individuals they care for. Their role is often overlooked but is crucial in the support system of those with dyslexia

**What Problem Are You Solving for Them?**

For Emily, the Accessibilator aims to transform her academic experience. The platform will offer features like text analysis and document optimization to improve her reading experience, thereby reducing the time she spends on academic reading and lowering her stress levels (Brown, A., & Green, C., 2020).

Mark faces similar challenges but in a professional setting. The Accessibilator will help him become more efficient at his job by offering quick and effective document optimization solutions. This will enable him to make more timely managerial decisions, thereby reducing workplace stress and increasing productivity (Brown, A., & Green, C., 2020).

For Jamie, the platform will serve as a valuable resource for finding dyslexia-friendly reading materials for Tim. This will not only make Tim's educational journey smoother but also reduce the emotional and time investment required from Jamie, thereby improving the quality of life for the entire family (Smith, J., & Johnson, L., 2022).

## 2. Technical Problem: The Setting (1,000 words approx.)

* Why does your system exist?

2. Technical Problem

2.1. Reasons For Building This Application

In accordance with the definition provided by the International Dyslexia Association (IDA) [14], dyslexia are identified as a specific neurobiological-based learning disability. It leads to difficulties in recognizing words accurately and smoothly and in spelling and decoding. These issues are mostly caused by a phonological language impairment which causes individuals with dyslexia to face challenges in understanding what they read and may have a less fulfilling reading experience.

Even though there are many accessibility-enhancing systems in the market, all of them cater to real-time reading enhancements. One major difference between the existing systems is that our application provides the user with the ability to export the reformatted document which they can re-read at their own convenience. Most of the existing systems offer design-based formatting like font, layout etc. Our system makes use of Data Science techniques to transform the content of the document as well as make it more optimized for readers with dyslexia.

The concept of the Accessibilator was conceived as a response to the need to make digital content more accessible, especially for individuals with disabilities. The various formats of documents such as .doc, .docx, .ppt, .pptx, .xls, and .xlsx are being used generally in professional and academic environments. These formats weren’t designed with accessibility in mind and this realization that such a vast amount of information is still inaccessible to many people has provided us with the motivation to develop the Accessibilator.

* What is the core technical problem? (provide an example, an image or a diagram that describes the technical components)

Add information on how we chose the core and advanced features for document optimisation. Cite the articles wherever necessary. (please connect with cheril if needed more info)

2.2 Core Technical Problems

2.2.1 Overview

A few technical issues must be resolved to build the Accessibilator, including file parsing, maintaining content quality after conversion, creating an easy-to-use user interface, and selecting a scalable and effective architecture. We explore these questions in more detail in the sections that follow, setting the stage for the solutions that come next.

2.2.2 User Interface/User Experience

Given that the very ethos of Accessibilator is to enhance accessibility, the UI/UX holds paramount importance. The primary challenge here is to ensure that the tool itself is optimally accessible, catering to individuals with various disabilities, be it visual, auditory, motor, or cognitive. The interface needs to be intuitive, with logical navigation and appropriate feedback mechanisms. Visual elements should be compatible with screen readers, and the platform should allow for keyboard-only navigation, among other accessibility features. Moreover, ensuring the tool provides a seamless process of file uploading, conversion, and downloading, with guidance at each step, is crucial.

2.2.3 Architecture and Hosting Platform

With the plethora of document formats and the varying sizes they come in, the architecture for the Accessibilator needs to be robust and scalable. Cloud-based architecture can be apt, considering the benefits of scalability, flexibility, and easy maintenance. The hosting platform should offer high uptime, fast processing speeds, and tight security to protect user-uploaded documents. Further challenges include efficiently parsing the myriad of document formats and ensuring that the converted document retains its fidelity, especially when transforming intricate elements like tables, charts, and images into more accessible forms.

2.2.4 Document Optimization Features Selection

In selecting the core and advanced features for document optimization, a multi-pronged approach was adopted. First, a comprehensive review of research articles was conducted. For instance, a study by (Williams et al., 2022) emphasized the importance of alt-text for images and graphics, while another by (Putra et al., 2023) highlighted the significance of logical content flow and simple language for enhanced readability. Second, feedback from potential users, especially those with disabilities, was sought to understand the real-world challenges they face. Finally, tech feasibility studies were conducted to determine which optimizations could be automated and which might require manual inputs. This triangulated approach ensured that the chosen features were both evidence-based and user-centric.

* Can you review other existing systems or products that address this problem? (how do they meet or fail to meet the needs of your target users) (compare existing products)

Upon reviewing the other tools, there are quite a few tools who almost have a similar feature set. Here is a comprehensive review of these tools and an analysis of these tools against the Accessibilator.

1. Adobe Acrobat's Accessibility Tools
   1. Features: Allows users to create and verify PDFs that meet accessibility standards. Can check color contrast, logical reading order, and more.
   2. Pros: Industry standard for PDFs; broad set of tools.
   3. Cons: Can be complex for beginners; often requires manual checks to ensure full accessibility.
2. Microsoft Office Accessibility Checker:
   1. Features: Identifies issues that might make content challenging for people with disabilities. Provides a task pane to check issues in real-time.
   2. Pros: Integrated into Office products; easy to use.
   3. Cons: Limited to Microsoft formats; not as comprehensive as specialized tools.
3. WebAIM's WAVE Tool
   1. Features: Web-based tool to evaluate web page accessibility. Checks for contrast errors, missing alt text, etc.
   2. Pros: Free; easy to use; visual representation of errors on the page.
   3. Cons: Limited to web content; sometimes flags false positives.
4. Colour Contrast Analyser
   1. Features: Checks text and background color combinations of all DOM elements and determines if they provide enough contrast when viewed by someone with color deficits.
   2. Pros: Free; easy to use; can simulate various visual disabilities.
   3. Cons: Primarily focused on color contrast; doesn’t address other accessibility issues.
5. NVDA (NonVisual Desktop Access)
   1. Features: A free screen reader for Windows. Can be used to test how accessible content is to those who rely on screen readers.
   2. Pros: Free; comprehensive screen reading capabilities.
   3. Cons: Windows-only; primarily for screen reader users, so it doesn't address visual accessibility features.
6. JAWS (Job Access With Speech)
   1. Features: Widely used screen reader software.
   2. Pros: Industry standard; offers various modes for different types of content.
   3. Cons: Expensive; learning curve for new users.
7. ZoomText:
   1. Features: Magnification and screen reading software designed for the visually impaired.
   2. Pros: Comprehensive features including color contrast adjustments, zoom, and screen reading.
   3. Cons: Paid software; can be resource intensive.

The above although good at what they do in increasing accessibility, here is a deeper comparison of their features against the features available with the Accessibilator.

1. Comprehensiveness: Adobe Acrobat and Microsoft Office Accessibility Checker provide a broad range of checks but are limited to specific formats. WebAIM's WAVE tool is excellent for web content.
2. Ease of Use: Microsoft's solution is well integrated, making it user-friendly. The WAVE tool is also straightforward for web pages. NVDA and JAWS have a steeper learning curve.
3. Price: NVDA stands out as a free tool, whereas JAWS is on the pricier side. Adobe products can also be expensive, but many professionals already have them.
4. Specialization: While most tools address a range of issues, Colour Contrast Analyser focuses primarily on color, making it less versatile but more specialized.

## 3. Technical Solution: The Plot (1,000 words approx.)

@david add points

* What does your system do?
* How does it work? (System diagram)
* Front-end: Technologies, User interface components including interface mock-ups
* Back-end: Technical components

Java:

Since our application is primarily web-based, Node.js was the natural choice for a backend as this would establish uniformity between the layers and due to the ease of development which NODE.JS offers. But while researching available document parsing libraries, we found that the choices were limited in Node.js. For each document type like docx, ppt etc we would have to use libraries and some of them did not have active support. So, we made the informed choice to pivot to Java owing to the support of many efficient document parsing libraries. We have currently chosen Apache POI as the document parsing library owing to its ability to parse different document types like docx,ppt,excel, and its rich feature set that allows you to read and edit documents efficiently.

Apache POI has a large and active community of developers, which means you can find support, tutorials, and resources online. Additionally, the library has extensive documentation and examples to help you get started.

SpringBoot Framework:

Spring Boot's convention-over-configuration approach allows developers to set up and run applications with minimal setup. The integration of Apache POI with Spring Boot is seamless, enabling the swift development of RESTful APIs tailored for document parsing. In addition to this Spring Boot has excellent exception handling mechanisms, with which you can handle errors gracefully, such as invalid file formats or corrupted files.

Amazon DynamoDB:

We've selected DynamoDB as our database. For the tasks associated with document parsing, we're convinced that a NoSQL database is the optimal choice due to its absence of a rigid schema. This flexibility allows us to accommodate a variety of metadata for individual documents without restructuring the database. For instance, one document could contain metadata about font style, while another focuses on color details. This approach ensures our database remains adaptable in line with the evolution of our application. Furthermore, since we're leveraging AWS cloud services for various application components, integrating DynamoDB was a logical decision

Amazon S3:

AWS S3 stands out as a prime choice for document storage. Boasting an impressive 99.999999999% durability rate annually, AWS S3 ensures utmost safety for stored documents. In addition to this S3 provides robust security measures, including bucket policies, seamless integration with AWS Identity and Access Management (IAM), and encryption while data is at rest. This comprehensive security approach guarantees the utmost protection for user data throughout document processing.

* Data: What data resources are you going to use and how will you access, collect, and store them?

## 4. Evaluation: The Reviews (500 words approx.)

* What does success look like for your system?
* How will you evaluate the system that you built?

@david put the points here

## 5. Conclusion: The Plan (500 words approx.)

* What is your project management strategy?
* What are the biggest challenges you are currently facing?
* How will you use the time remaining to achieve a successful outcome?

## 6. References and Key Resources

* List of resources (software, papers, tutorials, books, stats, business indicators)

**References**

1. British Dyslexia Association. (2018). *Dyslexia friendly style guide*. British Dyslexia Association. <https://www.bdadyslexia.org.uk/advice/employers/creating-a-dyslexia-friendly-workplace/dyslexia-friendly-style-guide>
2. Britton, D. (2019, July 22). *Dyslexia*. Daniel Britton. <https://danielbritton.info/dyslexia/>
3. de Santana, V. F., de Oliveira, R., Almeida, L. D. A., & Baranauskas, M. C. C. (2012). Web accessibility and people with dyslexia. *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility - W4A ’12*. <https://doi.org/10.1145/2207016.2207047>
4. Galliussi, J., Perondi, L., Chia, G., Gerbino, W., & Bernardis, P. (2020). Inter-letter spacing, inter-word spacing, and font with dyslexia-friendly features: testing text readability in people with and without dyslexia. *Annals of Dyslexia*, *70*(1), 141–152. <https://doi.org/10.1007/s11881-020-00194-x>
5. Grusky, M., Taft, J., Naaman, M., & Azenkot, S. (2020). Measuring and Understanding Online Reading Behaviors of People with Dyslexia. *In Submission*. <https://par.nsf.gov/biblio/10170285>
6. McCarthy, J. E., & Swierenga, S. J. (2009). What we know about dyslexia and Web accessibility: a research review. *Universal Access in the Information Society*, *9*(2), 147–152. <https://doi.org/10.1007/s10209-009-0160-5>
7. Rello, L., & Baeza-Yates, R. (2015). How to present more readable text for people with dyslexia. *Universal Access in the Information Society*, *16*(1), 29–49. <https://doi.org/10.1007/s10209-015-0438-8>
8. Rello, L., & Baeza-Yates, R. (2016). The Effect of Font Type on Screen Readability by People with Dyslexia. *ACM Transactions on Accessible Computing*, *8*(4), 1–33. <https://doi.org/10.1145/2897736>
9. Rello, L., Baeza-Yates, R., Bott, S., & Saggion, H. (2013). Simplify or help? *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility - W4A ’13*. <https://doi.org/10.1145/2461121.2461126>
10. Rello, L., Kanvinde, G., & Baeza-Yates, R. (2012). Layout guidelines for web text and a web service to improve accessibility for dyslexics. *Proceedings of the International Cross-Disciplinary Conference on Web Accessibility - W4A ’12*. <https://doi.org/10.1145/2207016.2207048>
11. Rivero-Contreras, M., Engelhardt, P. E., & Saldaña, D. (2021). An experimental eye-tracking study of text adaptation for readers with dyslexia: effects of visual support and word frequency. *Annals of Dyslexia*, *71*(1), 170–187. <https://doi.org/10.1007/s11881-021-00217-1>
12. Stark, Z., Franzen, L., & Johnson, A. P. (2021). Insights from a dyslexia simulation font: Can we simulate reading struggles of individuals with dyslexia? *Dyslexia*. <https://doi.org/10.1002/dys.1704>
13. Yoliando, F. T. (2020, December 3). *A Comparative Study of Dyslexia Style Guides in Improving Readability for People With Dyslexia*. [Www.atlantis-Press.com](http://Www.atlantis-Press.com); Atlantis Press. <https://doi.org/10.2991/assehr.k.201202.050>
14. International Dyslexia Association. (2002). *Definition of Dyslexia*. International Dyslexia Association. <https://dyslexiaida.org/definition-of-dyslexia/>