

Towards More Efficient Screen Reader Web Access with Automatic Summary Generation and Text Tagging

Usama Sarwar¹ and Evelyn Eika¹

¹ Oslo Metropolitan University, 0130 Oslo, Norway
s310272@oslomet.no, Evelyn.eika@oslomet.no

Abstract. Readers with 20/20 vision can easily read text and quickly perceive information to get an overview of the information within a text. This is more challenging for readers who rely on screen readers. This study investigated factors affecting successful screen reading in order to shed light on what contributes towards the improvement of screen reading access. Text extraction, summarization, and representation techniques were explored. The goal of this work leads to the development of a new summarization technique, referred to as On Demand Summary Generation and Text Tagging (ODSG&TT). This technique makes use of a summarization algorithm and a text tagging algorithm developed by Algorithmia, which enables on the fly and on-demand summarization of text and keyword generation. The focus of the screen reader is transferred to the keywords using a button control. The intention is to provide summaries with minimum user navigation effort to simplify the screen reading process.

Keywords: automatic summarization, screen reader, visually impaired, reading, skimming, accessibility.

1 Introduction

Digital information is increasingly being used in communications. This trend requires that citizens obtain information electronically. Hence digital information must be made easily available for all readers, including the visually impaired who rely on screen readers to gain access to information. Most screen readers that assist the visually impaired read the text in an iterative manner. Since these tools present text sequentially, it is difficult for users to scan for items of interest or importance compared to people who read text with their eyes. Browsing through the digital documents thus becomes a slow and often pains-taking process for the screen reader user. A more intelligent screen reader that leads to more efficient navigation through the electronic document would hold potential for greatly assisting users in accessing information. This study investigates challenges that screen reader users commonly encounter and explores text extraction techniques as a potential remedy for achieving more effective screen reading.

Screen readers are software tools developed to enable the visually impaired to use computers. These tools read aloud with a synthetic voice and they are commonly used by the visually impaired to access the web. Readers who rely on their eyes form an overview of the contents by simply visually scanning the webpage; the screen reader

users are not using their eyes but have to construct what is there bit by bit [1]. What visually oriented readers regard as a simple task such as getting advice online for a web product/service may not be trivial for a screen reader user [1]. As expressed by a user who had limited experience with screen readers, intensive navigating through the library's websites took much longer time (twice or three times as much) to search for information compared to visually oriented readers [2]. Though screen readers assist users greatly, limitations exist and need to be resolved to provide more accessible use. Some weaknesses include reading irrelevant visual presentation elements, employing a simple top-to-bottom and left-to-right reading strategy leading to lengthy waiting for relevant information, unable to deliver structural information of the page, intertwining content reading with links reading that causes confusions and inconveniences for listeners who have to wait for links' meanings to be clarified or select long wrong moves in websites due to ineffective ordering of links' list, impractical links selection scheme because of audio synchronization difficulties, along with the synthetic voice reading all pieces of information on the page (content, links, landmarks, etc.) with the same intonation throughout [3].

Automatic text summarization techniques are proposed as a means to help screen reader users reduce cognitive workload by generating meaningful summaries of long texts. Many artificial Intelligence algorithms are developed based on natural language processing. However, the process of using such software is not simple. It often involves navigation challenges for screen reader users as they have to shift their focus from different applications. Summarization may include multiple steps: (1) Select the text to summarize. (2) Copy the selected text. (3) Open the text summarizer through webpage, plugin link, or desktop software. (4) Paste the copied text into the software. (5) Generate the summary. (6) Read it through the screen reader. (7) Come back to the original text (webpage or other electronic text). (8) Users must start reading again from the initial point as the context is lost. These steps may seem simple but require effort; the overall purpose of saving effort is thus lost. Moreover, some web-based summary generators return the summaries in dialog boxes or alerts, and such popups reduce accessibility.

Given the challenges that users encounter when navigating using the screen reader, this work aims to minimize the efforts involved through an integrated approach where the summary is available as part of the regular screen reader reading flow. The following questions are asked: Can effective screen reading be achieved with minimized navigation? Can lengthy text reading of website screen reading be simplified?

This paper is organized as follows. Related work is presented, followed by sections of method and results. Discussion is then offered based on findings gathered, closed by a conclusion.

2 Related work

According to Borodin et al. [4], browsing the Web with screen readers can be challenging because of accessibility and usability problems. Accessibility issues associated with screen reader users have also been identified in websites of higher education [5]. The basic readability of the text affects all readers [6-9], and it has been demonstrated that

readability features such as sentence length affect screen reading performance [10]. It is nearly impossible to access some content through screen readers. One common example is an image without an alternative text. Screen readers will ignore such images which then become “invisible” to the readers. Though the screen reading technology is advancing with new features, new web technologies emerge at an even faster pace and the discrepancies lead to inaccessibility. Implementing an accessible website without any rich media is by some viewed as laborious [11] as W3C recommendations are implemented differently in different browsers. Ensuring accessibility with rich media is even more challenging. For instance, the Flash multimedia content was inaccessible to screen reader users for years [4]. With HTML 5 introducing canvas, content inaccessibility may arise again that will affect the working of web technologies with screen readers.

Screen readers are usually slow to use because they read out the content sequentially at a fixed rate. It is known that interactions of screen reader users are slower than those of users who make use of the visual interface [11, 12]. Repeated navigation links and complex search functions were among main accessibility issues for screen reader users [13]. As also proposed, users’ needs such as adapting the speed of the screen reader should be considered in web accessibility evaluations [14]. Still, navigational constraints remained in current screen readers, and multimodal technologies enhancing non-visual interface were proposed to provide navigational cues to assist users [15].

Skimming is an effective way of quickly reviewing large amount of information, where the readers’ eyes rapidly explore the visual contents. However, this is not easy for screen reader users. Screen readers typically rely on text-to-speech techniques or tactile braille displays to convey the digital content non-visually allowing users to navigate through the content marked up by HTML tags for headings, paragraphs, links, and buttons [16]. A screen reader scans the whole page from start till the end that makes the process sequential and potentially time-consuming. Screen-readers have features to accelerate access where users can jump more rapidly through the content a sentence at a time, jump across lines, paragraphs, sections, pages, and traverse lists of links and headings. Although these functions provide great improvements, they do not provide the same flexibility, speed, and freedom as visual skimming of information would [17].

Although as a comprehension strategy the screen reader user could filter through the document by choosing a faster screen reading speed, it has drawbacks of increasing users’ cognitive load since more concentration is required. The increased attention thus demanded could reduce word comprehension as some participants experienced and result in overall reduced comprehension [18]. Further, linear presentation gives users little room for maneuvering information selection hence fatiguing their working memory, which in turn is found to reduce efficiency of information processing [19]. There exists, also, a trade-off between speed reading and comprehension accuracy, and fast reading pace that hinders reread could cause misinterpretation because of word ambiguity and sentence complexity [20]. As recommended, screen reader users should be given flexibility to change reading speed as current tools already offered, or to relocate information space for individual users’ preferences [18].

Because of the low bandwidth of a serial audio interface and braille displays, screen reader users spend considerably more time identifying the information they need compared to users who skim visually [4]. The cognitive load increases when a user has to process a large amount of information [21]; for screen reader users, the cognitive load is even higher as they have to listen, remember, and process the content that is presented sequentially. As experienced by some users, they resort to listen to the whole document since it is nearly impossible to identify specific parts of the document [18]. The assumption of this work is that by reducing the amount of information through summarizing techniques consequently reduces the cognitive load of screen reader users.

Automatic summarization approaches are used to summarize the textual content while preserving its essence [22]. Two main classes of text summarization techniques exist, namely summarization by extraction and summarization by abstraction [22]. According to [23], summarization by extraction deals with selecting the most prominent and meaningful phrases from the content. Summarization by abstraction is constricting sentences [22, 24], which involves rephrasing the content by automatically filling out a template [25].

Various algorithms have been proposed to foster faster screen reading. For example, skim reading is available in JAWS (Job Access With Speech) 6.0 and later. Skim reading refers to gaining the gist of a text without using all the details [20]. This feature has a simple implementation where users can swiftly review long documents by reading the first line or sentence of each paragraph. JAWS also provides users with looking up for specific phrase and words and with generating summaries of the text for users' overview of the whole content and for detailed reading if it is relevant. Users could also make rules to read in certain patterns, but they cannot be applied to all contents. Users' experience was that JAWS skimming was not useful for content skimming and it was similar to navigating paragraphs with a different purpose, hence not efficient [16, 18]. Specifically, users indicated that they used regular navigation shortcuts for skimming since no tools provided adequate content skimming support [16].

Accessible skimming is a concept introduced by Ahmed et al. [16]; it is also a form of non-visual skimming. The summarization algorithm uses a text extraction technique with the following steps: (1) Generate variable sized summaries of the text. (2) Each sentence is summarized to cover the complete text. (3) Phrases are extracted such that the meaningful connections between the words hold. This technique has been successful in facilitating screen reading as opposed to using regular shortcuts [16]. An additional shortcut was enabled to let users switch between the summary and the original text while preserving the reading position within the text so that the users could return to the same spot after listening to the generated summary without having to reread from the beginning of the article [16].

Tag Thunder [26] is a cloud content representation system. It works based on principles of content summarization and concurrent speech synthesis [27]. Tag Thunder implements content presentation as follows: The Web Page is segmented into several zones, then key words are extracted through text extraction techniques, and extracted key terms are concurrently vocalized on an audio track to echo the positions and visual properties of respective zones. It enables users to filter out the content of interest and determine the desired zone for further navigations.

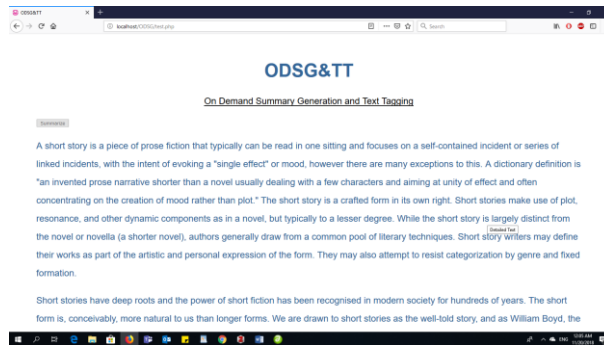


Fig. 1. ODSG&TT enabled webpage.

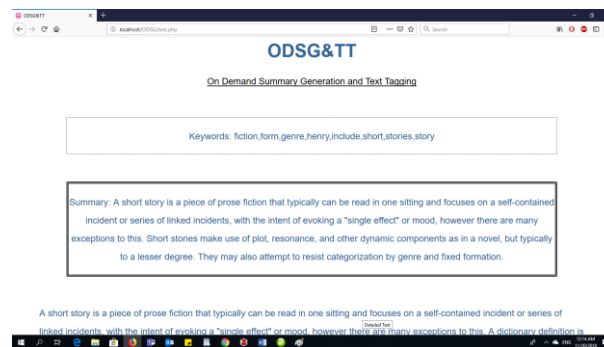


Fig. 2. The webpage appears with ODSG&TT executed.

3 Method

To answer the research questions, an experiment with a prototype involving screen reader users and questionnaire feedback was conducted. The goal was to contribute towards improved screen reading.

3.1 Equipment

Nonvisual Desktop Access (NVDA) screen reader was used for the experiment. NVDA ensures nonvisual access to and interaction with Windows OS and other applications [28].

3.2 Prototype framework

A prototype named On Demand Summary Generation and Text Tagging (ODSG&TT) was implemented using PHP version 5.6.35 and the jQuery library version 3.3.1. Because this was a controlled experiment, we only used localhost with a WAMP server. WampServer is a Windows web development environment [29]. We used the Summarizer summarizing algorithm [30] developed on Algorithmia [31] platform. It is a developed algorithm and can be used with an API Key. AutoTag [32] is another library developed at Algorithmia [31] that was used to extract keywords of texts. The prototype setup was run on a computer running Windows 10 Enterprise version with Mozilla Firefox version 63.0.3 (64-bit) and Google Chrome Version 70.0.3538.102 (Official Build, 64-bit).

Fig. 1 shows a website with the ODSG&TT functionality enabled. Users can generate the keywords and summary of the text within the webpage; they can also decide if they want to skip the keywords and summary and continue.

Fig. 2 shows that there are two additional sections generated within the same webpage shown in Fig. 1, namely the sections keywords and summary. The original text is still there, and it will be read after the screen reader reads out these two sections. This gives users a rapid clue about whether the text is relevant and interesting or not. If yes, the screen reader will continue reading the detailed text as well; otherwise, the user can simply move to another webpage.

3.3 Questionnaire

An interview using five questions was conducted to elicit feedback. The questions addressed the efficiency of ODSG&TT: Does it serve the purpose of summary generation? Is it faster than the other two summary generation techniques (see the procedure section)? Is it simple to use? Does it address the navigation issues? Does it simplify the screen reading when generating summaries?

3.4 Participants

Twenty participants ($N = 20$) were recruited for the study. They all had information technology backgrounds and knew how to use screen readers or were familiar with concepts of using screen readers.

3.5 Procedure

The participants were first given instructions about the task. Participants were blindfolded as it was challenging to recruit actual screen reader users. One-to-one sessions were conducted. A summary was generated from a website and read with the NVDA screen reader; the time taken was measured using Resoomer. A second summary was generated using a browser plugin and read using the NVDA screen reader; the time

taken was measured via Chrome Plugin Text-Summarizer. A third summary was generated using ODSG&TT (also read with the NVDA screen reader), and the time taken was measured. A questionnaire-based interview was then conducted.

3.6 Analysis

A one-way repeated ANOVA was used, with summarization technique as the independent variable and summarization time as the dependent variable. SPSS was used for the statistical computations.

4 Results

4.1 Processing time of summarization techniques

The results indicated that compared to the Plugin and Website processes, the ODSG&TT summarization technique required less time. These repeated measures were conducted at a significance level .05 with a 95% confidence interval. The following figure indicates the mean time taken in seconds to generate the summary of the text.

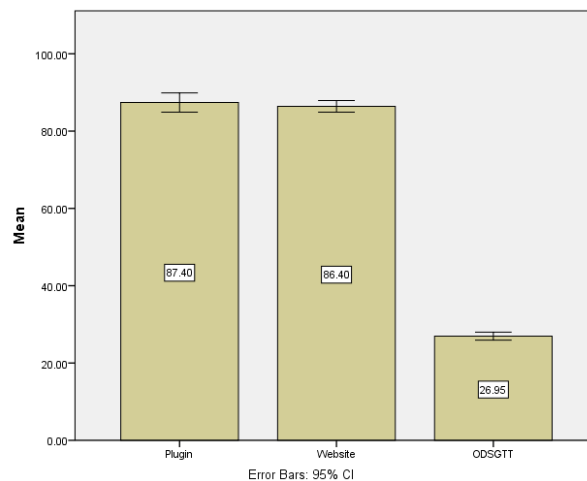


Fig. 3. Bar Chart for Calculated Means with 95% Confidence Interval.

The observations did not violate the assumption of sphericity ($\chi^2(2) = 3.718, p = .156$). The ANOVA showed that there was a significant difference in summarization time for the three summarization techniques ($F(2, 38) = 1439.5, p < .001, \eta p^2 = .987$). Bonferroni post-hoc analysis revealed that the plugin summarization technique ($M = 87.4, SD = 5.315$) and the website summarization technique ($M = 86.4, SD = 3.22$) exhibited almost the same summary generation time. Hence, the two techniques were not significantly different. However, the ODSG&TT summarization technique ($M =$

26.9, $SD = 2.23$) took significantly less time. These results support the claim that ODSG&TT is a faster summarization technique compared to the two other alternatives.

4.2 Questionnaire feedback

Most participants (80%) expressed that ODSG&TT addressed the purpose satisfactorily and succeeded in performing the intended functionality. However, one-fifth of the participants (20%) thought it did not entirely address the purpose. In response to the second question, as high as 95% of the participants perceived that the prototype was faster than the two other summary generation techniques while only 5% of the participants did not perceive that it was any faster. As for simplicity, most of the participants found the prototype simple and straight forward to use as no complex calculations or processes were involved. More than half of the participants (70%) responded that this system was simple while 30% thought it was moderately simple.

Concerning navigation challenges, most (75%) were of the opinion that ODSG&TT addressed the navigation challenges while one-fourth of the participants (25%) disagreed. As for simplifying the screen reading experience, the majority of the participants (85%) responded that it did simplify the screen reading while 15% disagreed.

5 Discussion

This research aimed to reduce the navigation challenges of screen reader technology by exploring a new summarization technique.

Perceiving and understanding information quickly is not straightforward with screen readers. Navigation challenges within digital content indicate that improved techniques and approaches are needed to achieve faster and effective screen reading with less cumbersome navigation operations. This work shows how the access to lengthy text on the web through screen readers can be simplified.

Existing summarization techniques require users to perform many steps. In order to minimize the number of steps, we developed a prototype named On Demand Summary Generation and Text Tagging (ODSG&TT). This prototype uses the existing summarization and text tagging algorithms and incorporates them within the website. The prototype helps screen reader users summarize long text using a button control. By pressing the button the summarized text and keywords are dynamically displayed within the webpage, and the focus of screen reader is also transferred to the keywords section.

With the current implementation, the regular text reading starts automatically after reading the summary. Some participants also suggested that there should be a pause with instruction that lets users know that summary is finished and if they want to continue or change the search. Future work may enhance the prototype by combining with smart devices to provide users with flexibility. Users may be able to give a voice command to devices such as Alexa and Siri to generate the summary of an article or text, and the prototype can run in the backend. A summary will then be generated and read aloud by the device.

Overall, the results show that the proposed technique does simplify the screen reading process because the focus of the screen remains within the webpage and the navigation procedures are reduced. The screen reader users do not have to swap the focus between different applications to access content.

6 Conclusion

Visual reading is not the same as screen reading. Screen reader users face several navigation challenges. The work shows that a lengthy text can be summarized within the same webpage, eliminating the need to shift focus between a given webpage and a summarizing tool. After reading the summary users can decide whether or not they want to continue. Participants' feedback also shows that ODSG&TT does not involve complex navigation and control transfers, making it faster and easier to use. Future work involves enhancing the prototype by letting users know when a summary is finished and explicitly asking them if they want to continue reading.

References

1. Gerber, E.: The benefits of and barriers to computer use for individuals who are visually impaired. *Journal of Visual Impairment & Blindness* **97**(9), 536-550 (2003).
2. Mulliken, A.: There is Nothing Inherently Mysterious about Assistive Technology: A Qualitative Study about Blind User Experiences in US Academic Libraries. *Reference and User Services Quarterly* **57**(2), 115-126 (2017).
3. Di Blas, N., Paolini, P., Speroni M., Capodiecchi, A.: Enhancing accessibility for visually impaired users: the Munch's exhibition. In: Bearman D. & Trant J. (eds), *Museums and the Web 2004. Selected Papers from an International Conference, Archives & Museum Informatics, Arlington, Washington, U.S.A., (2004)*.
4. Borodin, Y., Bigham, J. P., Dausch, G., Ramakrishnan, I. V.: More Than Meets the Eye: A Survey of Screen-reader Browsing Strategies. *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility* (2010).
5. Parajuli, P., Eika, E.: A Comparative Study of Accessibility and Usability of Norwegian University Websites for Screen Reader Users Based on User Experience and Automated Assessment. In: *International Conference on Universal Access in Human-Computer Interaction. LNCS vol. 12188, Cham: Springer* (2020). Doi: 10.1007/978-3-030-49282-3_21
6. Eika, E., Sandnes, F. E.: Assessing the Reading Level of Web Texts for WCAG2.0 Compliance—Can It Be Done Automatically? In: *Advances in Design for Inclusion*, pp. 361-371, Springer (2016).
7. Eika, E., Sandnes, F. E.: Authoring WCAG2.0-compliant texts for the web through text readability visualization. In: *International Conference on Universal Access in Human-Computer Interaction*, pp. 49-58. LNCS, Vol. 9737, Springer, Cham (2016). Doi: 10.1007/978-3-319-40250-5_5
8. Eika, E.: Universally designed text on the web: towards readability criteria based on anti-patterns. *Stud. Health Technol. Inform* **229**, 461-470 (2016).
9. Kaushik, H. M., Eika, E., Sandnes, F. E.: Towards Universal Accessibility on the Web: Do Grammar Checking Tools Improve Text Readability? In: *International Conference on Universal Access in Human-Computer Interaction. LNCS, Vol. 12188, Cham: Springer* (2020). Doi: 10.1007/978-3-030-49282-3_19

10. Kadayat, B. B., Eika, E.: Impact of Sentence length on the Readability of Web for Screen Reader Users. In: International Conference on Universal Access in Human-Computer Interaction. LNCS, Vol. 12188, Cham: Springer (2020). Doi: 10.1007/978-3-030-49282-3_18
11. Chapdelaine, C., Gagnon, L.: Accessible Video description On-Demand. In: Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility, 221-222 (2009). doi: 10.1145/1639642.1639685
12. Sun, S.-Y., Shieh, C.-J., Huang, K.-P.: A research on comprehension differences between print and screen reading. *South African Journal of Economic and Management Sciences*, 16(5), 87-101 (2013).
13. Gerber, E.: Technology and persons with disabilities. In: Proceedings of the 17th Annual International Conference of California State University Northridge (CSUN). (2002).
14. Vigo, M., Kobsa, A., Arrue, M., Abascal, J.: User-tailored web accessibility evaluations. In: Proceedings of the eighteenth conference on Hypertext and hypermedia, pp. 95-104 (2007).
15. Murphy, E., Kuber, R., McAllister, G., Strain, P., Yu, W.: An empirical investigation into the difficulties experienced by visually impaired Internet users. *Universal Access in the Information Society* 7(1-2), 79-91 (2008).
16. Ahmed, F., Borodin, Y., Soviak, A., Islam, M., Ramakrishnan, I. V., Hedgpeth, T.: Accessible Skimming: Faster Screen Reading of Web Pages. In: Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, pp. 367-378, ACM (2012). doi: 10.1145/2380116.2380164
17. Bigham, J. P., Cavender, A. C.: Evaluating Existing Audio CAPTCHAs and an Interface Optimized for Non-visual Use. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 1829-1838, ACM (2009). <https://doi.org/10.1145/1518701.1518983>
18. Machulla, T., Avila, M., Wozniak, P., Montag, D., Schmidt, A.: Skim-reading Strategies in Sighted and Visually-impaired Individuals: A Comparative Study. In: Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference, pp. 170-177, (2018).
19. Wickens, C. D., Gordon, S. E., Liu, Y.: An introduction to human factors engineering. Course material (1998).
20. Rayner, K., Schotter, E. R., Masson, M. E., Potter, M. C., Treiman, R.: So much to read, so little time: How do we read, and can speed reading help? *Psychological Science in the Public Interest* 17(1), 4-34 (2016).
21. Sweller, J.: Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12(2), 257-285 (1988). [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
22. Ahmed, F., Borodin, Y., Puzis, Y., Ramakrishnan, I. V.: Why read if you can skim: towards enabling faster screen reading. In: Proceedings of the International Cross-Disciplinary Conference on Web Accessibility, pp. 1-10, (2012).
23. Goldstein, J., Kantrowitz, M., Mittal, V., Carbonell, J.: Summarizing Text Documents: Sentence Selection and Evaluation Metrics. In: Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 121-128, ACM (1999). <https://doi.org/10.1145/312624.312665>
24. Berger, A. L., Mittal, V. O.: OCELOT: A System for Summarizing Web Pages. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 144-151, (2000). <https://doi.org/10.1145/345508.345565>
25. Jacobs, P. S.: To Parse or Not to Parse: Relation-driven Text Skimming. In: Proceedings of the 13th Conference on Computational Linguistics - Volume 2, pp. 194-198, (1990). doi: 10.3115/997939.997973

26. Manishina, E., Lecarpentier, J.-M., Maurel, F., Ferrari, S., Busson, M.: Tag Thunder: Towards Non-Visual Web Page Skimming. In: Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility, 281-282, ACM (2016). doi: 10.1145/2982142.2982152
27. Guerreiro, J., Gonçalves, D.: Scanning for Digital Content: How Blind and Sighted People Perceive Concurrent Speech. *ACM Trans. Access. Comput.* **8**(1), (2016). doi: 10.1145/2822910
28. NVDA: About NVDA. Retrieved November 18, 2018, from NV Access website: <https://www.nvaccess.org/about-nvda/> (2017).
29. WampServer: Retrieved November 19, 2018, from WampServer website: <http://www.wampserver.com/en/> (n.d.).
30. Algorithmia: Summarizer - Algorithm by nlp. Retrieved November 19, 2018, from Algorithmia website: <https://algorithmia.com/algorithms/nlp/Summarizer> (n.d.).
31. Algorithmia: Algorithmia - Deploy AI at Scale. Retrieved November 19, 2018, from Algorithmia website: <https://algorithmia.com> (n.d.).
32. Algorithmia: AutoTag - Algorithm by nlp. Retrieved November 19, 2018, from Algorithmia website: <https://algorithmia.com/algorithms/nlp/AutoTag> (n.d.).