Enabling Convenient Online Collaborative Writing for Low Vision Screen Magnifier Users

Hae-Na Lee
Stony Brook University
Stony Brook, NY, USA
haenalee@cs.stonybrook.edu

Yash Prakash
Old Dominion University
Norfolk, VA, USA
yprak001@odu.edu

Mohan Sunkara
Old Dominion University
Norfolk, VA, USA
msunk001@odu.edu

IV Ramakrishnan
Stony Brook University
Stony Brook, NY, USA
ram@cs.stonybrook.edu

Vikas Ashok
Old Dominion University
Norfolk, VA, USA
vganjigu@odu.edu

ABSTRACT

Online collaborative editors have become increasingly prevalent in both professional and academic settings. However, little is known about how usable these editors are for low vision screen magnifier users, as existing research works have predominantly focused on blind screen reader users. An interview study revealed that it is arduous and frustrating for screen magnifier users to perform even the basic collaborative writing activities, such as addressing collaborators' comments and reviewing document changes. Specific interaction challenges underlying these issues included excessive panning, content occlusion, large empty space patches, and frequent loss of context. To address these challenges, we developed *MagDocs*, a browser extension that assists screen magnifier users in conveniently performing collaborative writing activities on the Google Docs web application. MagDocs is rooted in two ideas: *(i)* a custom *support interface* that users can instantly access on demand and interact with collaborative interface elements, such as comments or collaborator edits, *within* the current magnifier viewport; and *(ii) visual relationship preservation*, where collaborative elements and the corresponding text in the document are shown close to each other within the magnifier viewport to minimize context loss and panning effort. A study with 15 low vision users showed that MagDocs significantly improved the overall user satisfaction and interaction experience, while also substantially reduced the time and effort to perform typical collaborative writing tasks.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**; **Empirical studies in accessibility**.

KEYWORDS

Online Collaborative Writing, Low Vision, Screen Magnifier, Visual Impairment, Accessibility, Assistive Technology

HT '22, June 28-July 1, 2022, Barcelona, Spain

ACM ISBN 978-1-4503-9233-4/22/06…$15.00
https://doi.org/10.1145/3511095.3531274

ACM Reference Format:
Hae-Na Lee, Yash Prakash, Mohan Sunkara, IV Ramakrishnan, and Vikas Ashok. 2022. Enabling Convenient Online Collaborative Writing for Low Vision Screen Magnifier Users. In *Proceedings of the 33rd ACM Conference on Hypertext and Social Media (HT '22), June 28-July 1, 2022, Barcelona, Spain.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3511095.3531274

1 INTRODUCTION

The use of online collaborative writing applications has significantly increased in recent years, as these applications enable a group of individuals to not only conveniently and simultaneously edit documents, but also check each other's updates and comments [16, 18, 34]. Despite the growing importance of such applications in professional and academic settings [8, 23, 29, 31, 39], their usability with regard to low vision screen magnifier users remains unexplored, as existing research works have primarily focused on the usability issues and needs of blind screen reader users [11, 12, 37, 38].

A screen magnifier (e.g., ZoomText [45], Apple Zoom [22], Windows Magnifier [35]) is an assistive technology that enlarges all application content including text and graphics. However, content enlargement pushes many portions of the application off the screen, and therefore screen magnifier users have to frequently *pan* the enlarged content to access occluded portions. This *content occlusion* and the associated *excessive panning* have been previously found to cause many usability problems for people with low vision in general web browsing [3, 4, 49], and therefore they can also potentially cause similar issues in online collaborative writing tools for screen magnifier users. To uncover these usability issues, we conducted an interview study with low vision users having diverse eye conditions and visual acuity, who frequently interact with online collaborative editors using a screen magnifier.

Our findings from the interview study indicate that the current interface layouts of many collaborative tools are not suitable for convenient low vision interaction using a screen magnifier software, primarily due to incompatibility between the distributed visualization of GUI elements in these layouts and the narrow views provided by the screen magnifier. Nearly all participants in the interviews stated that they struggled to perform the two basic collaborative activities: *(i)* viewing and addressing collaborators' comments; and *(ii)* reviewing the list of prior edits or changes by collaborators. The notable reasons underlying these problems included occlusion-induced loss of visual connections that define relationships between

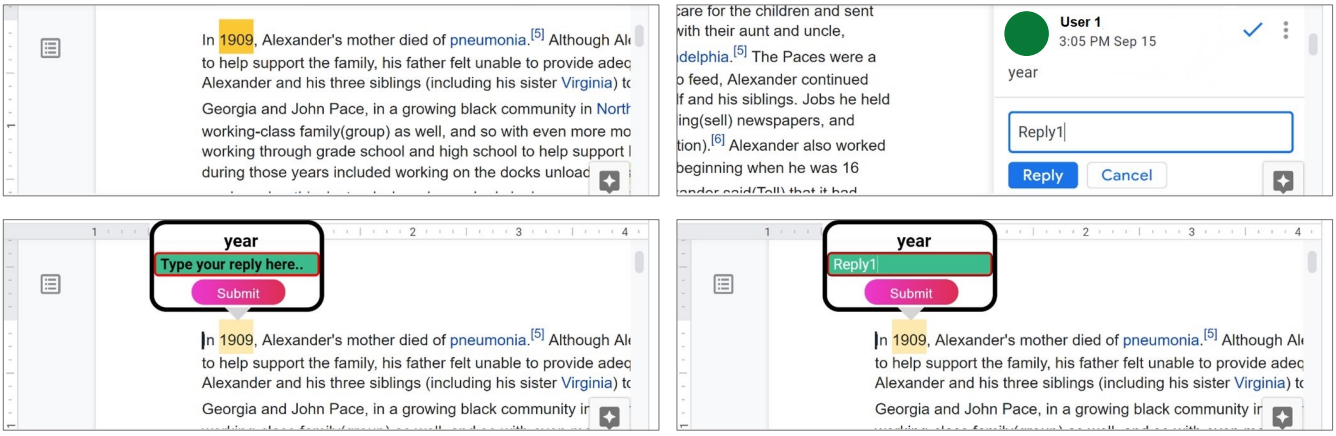


Figure 1: Comparison of use cases between a screen magnifier and MagDocs in Google Docs: (top row) – a low vision user accessing a comment in a document after magnifying screen content using the browser's built-in magnification feature. Notice that the user cannot view both the comment and the associated text at the same time within the same viewport; (bottom row) – a low vision user accessing a comment with MagDocs. In this case, both the comment and the corresponding text in the document are visible within the same viewport. The user can also respond to the comment in-place via the MagDocs popup.

different application segments (e.g., a comment and the corresponding text in a document), disorientation and loss of focus due to large empty whitespace patches while panning, and excessive to-and-fro panning between different application segments (e.g., document text and comments). Figure 1 (top row) illustrates one such usability problem faced by screen magnifier users while accessing collaborators' comments.

As an initial step towards addressing these usability concerns, we developed MagDocs, a browser extension for Google Docs, that enables low vision screen magnifier users to conveniently access collaborators' comments and document revisions with reduced panning effort. The two key ideas underlying MagDocs are: *(i)* a custom *support interface* – The viewport is automatically adjusted or re-positioned to bring the comments or changes one-by-one directly in front of the user, instead of user manually searching for them via panning; and *(ii) visual relationship preservation* – The comments and changes are shown right next to the associated text in the document to preserve the context, as opposed to the default application view where the comments and changes are shown to the far right on the screen. For example, Figure 1 (bottom row) illustrates how MagDocs assists screen magnifier users in accessing comments. To provide these functionalities to the users, MagDocs leverages custom built Mask R-CNN models that can automatically identify and extract comments and document changes respectively from the application document object model (DOM). Evaluation of MagDocs in a user study showed significant improvement in usability as well as access time overhead, compared to those with a state-of-the-art screen magnifier.

In sum, our contributions in this paper are as follows:

- Findings of an interview study uncovering usability issues faced by low vision users in online collaborative editors.
- Design and development of MagDocs, a browser extension for Google Docs that facilitates easy and instant access to comments and document changes, thereby mitigating the adverse effects of content occlusion and excessive panning.

2 RELATED WORK

Our research is relevant to extant literature on: *(i)* low vision usability issues, *(ii)* usability of online collaborative writing applications, and *(iii)* usability-enhancing solutions for low vision users.

2.1 General Low Vision Usability Issues

While there are a plethora of existing research works that have studied the usability issues of blind screen reader users [5, 7, 26, 33], usability issues faced by low vision users who rely on screen magnifier assistive technology have been largely underexplored [24, 36, 49]. The few existing works in this regard have mostly focused on uncovering the general usability concerns of low vision users. For example, Jacko et al. [24] focused on low vision users with age-related macular degeneration, and analyzed cursor movement control by these users. In their study, they observed that larger GUI element sizes yielded better user performances. A study by Szpiro et al. [49] investigated how people with low vision used mainstream computing devices for accessing information, and whether the current digital low vision accessibility tools provided adequate support in this regard. They found out that low vision users preferred visually accessing information more than relying on speech, and furthermore utilized multiple accessibility tools side-by-side to perform everyday tasks such as reading online news articles.

2.2 Usability of Online Collaboration Tools

Despite the importance of online collaborative tools, there is a dearth of studies that understand and address the concerns of people with visual impairments. Almost all existing works in this regard have focused on understanding the interaction challenges of blind screen reader users, and subsequently providing solutions to mitigate their problems [10, 13, 27, 38, 44]. For example, Mori et al. [38] proposed virtual overlays as a feasible solution to various screen reader-related usability issues. Similarly, Das et al. [13] conducted interviews with visually impaired people (all blind screen reader

users) who frequently used collaborative writing applications with sighted peers to understand their concerns. The interviews revealed that blind users typically had to go through arduous processes such as learning an ecosystem of collaborative tools, adapting to the complexities of collaborative features, adjusting the cost and benefits of accessibility, and learning to interact with sighted colleagues in groups. Other similar usability solutions for blind users have been proposed in recent years [27, 44, 50]. For instance, the CollabAlly system [27] provided collaborative and contextual information in a document such as active collaborators, comments, and changes, to screen reader users via audio features including spatial audio and voice fonts. Perhaps the closest related research to our work is by Lee et al. [28], who studied usability problems faced by low vision users on desktop productivity applications such as Microsoft Word. They also proposed a solution to address various issues, by enabling low vision users to quickly access ribbon menu controls via a proxy interface to facilitate a WYSIWYG (What You See Is What You Get) feedback and thereby reduce panning. However, the solution does not provide support for the collaborative aspects or features of the application such as comments and revisions.

2.3 Low Vision Usability-Enhancing Techniques

To the best of our knowledge, there are no prior works directly addressing the usability issues of low vision users with regard to online collaborative applications. Extant usability solutions for low vision users have mostly dealt with generic GUI enhancements [14, 15, 25, 32], web browsing [3, 4], and generic smartphone interaction [40]. Among these topics, web browsing for low vision users has especially received relatively more attention [1, 3, 4]. For instance, Bigham et al. [3] suggested the idea of opportunistic usability improvement and developed a magnification system for web browsing that automatically generated enlarged webpage content without adverse side effects, such as additional horizontal scrolling. Similarly, Billah et al. [4] designed a context-preserving screen magnifier equipped with a custom space compaction method, which ensured that all local and relevant web elements were kept close to each other within a magnified viewport. Although the prior solutions indeed help mitigate usability issues, they are presently unable to handle scenarios that require users to simultaneously view application segments that are spatially located far from each other. MagDocs was specifically designed to address such cases in the context of online collaborative writing applications.

3 UNCOVERING USABILITY ISSUES

We conducted an interview study with low vision screen magnifier users to understand their interaction challenges and needs with regard to online collaborative writing tools.

3.1 Participants

We recruited 10 low vision users through email lists and word of mouth. The average age of participants was 45.4 (Median = 46, SD = 11.8, Range = 31-62), and we ensured that the gender representation was approximately even (4 female, 6 male). The criteria for inclusion were: (a) strictly screen magnifier users, so that users with extremely poor visual acuity relying on screen readers were excluded from the interview; (b) proficiency in one of the following screen magnifiers – ZoomText, Apple Zoom, Windows Magnifier [51]; and (c) familiarity with online collaborative writing tools such as Google Docs. Table 1 (in Appendix A) presents the participant demographics. All participants were aware of their eye conditions, which encompassed a diverse range including retinitis pigmentosa, Stevens-Johnson syndrome, and glaucoma. The visual acuity of the participants ranged between 20/100 and 20/500. Also, only two participants stated that they sometimes used the speech narration feature of screen magnifiers. None of the participants had any motor impairments that affected their interaction with computer applications.

3.2 Interview Format

The interviews were all done remotely via either phone, Skype, or Zoom. The interviews had a semi-structured format, with the following questions:

- General questions about screen magnifiers and online collaborative writing tools: Which screen magnifier do you use? Which browser do you use? Which online collaborative tools do you use? For what purpose do you require online collaborative tools?
- Specific questions regarding their interaction experience with online collaborative writing tools: What activities do you typically perform on these tools? What application features do you access to perform these activities? Do you face any issues while doing these activities or while accessing a specific application feature?

Each interview started after obtaining the participant's consent and lasted about 45 minutes. The participants were compensated with an Amazon gift card for their time. The interview responses were transcribed and qualitatively analyzed using an open coding technique [43], where we iterated over the responses to capture recurring topics and usability issues. Some of the notable observations are presented next.

3.3 Findings

Use of collaborative tools driven by job requirements. All participants stated that they used an online collaborative editor, specifically Google Docs, in their profession. Most (8) participants also indicated that this was mainly due to peer pressure or 'team preference'. The remaining 2 participants stated that they preferred using Google Docs, as it was more convenient than exchanging several 'emails with attachments' with their peers.

Difficulty in associating comments with the corresponding text. All participants indicated that addressing peer comments was one of the main activities they performed while collaboratively editing documents. However, all participants stressed that this was a challenging task due to occlusion of content induced by screen enlargement. Specifically, the participants mentioned that in the current interface they cannot view both a comment and the corresponding text in a document within the same viewport of their screen magnifiers. Therefore, they have to manually *pan* the screen between the comment and its associated text in order to understand the underlying context and subsequently interpret the meaning of the peer comment. Six participants explained that this process was

stressful and frustrating because they often have to pan to-and-fro between the comment and the document text multiple times to fully understand the comment.

Difficulty in accessing and navigating over unresolved comments. Almost all (9) participants mentioned that it was both tedious and strenuous to sequentially navigate to different portions of the document in order to address different comments from collaborators one-by-one. They specifically mentioned that this activity involved considerable manual search effort and hence excessive panning, given that only a small portion of the screen was visible to them at any instant.

Difficulty in accessing and interacting with version history. All participants stated that they frequently reviewed changes in document content made by their collaborators. However, the participants listed similar usability issues as in case of comments, presumably due to the similarity in how Google Docs renders the list of comments and the list of revisions in its graphical user interface. For instance, all participants indicated that they expended significant amount of panning effort to associate a change in the document with not only the name of the collaborator making that change, but also the date and time of the change. Also, a majority (8) of the participants expressed that it was arduous to search and go over the list of changes one-by-one in the document.

4 APPROACH

Informed by the findings of the interview study, MagDocs specifically focuses on improving usability of two collaborative features in Google Docs – *comments* and *document changes*. These two features share a common trait of involving multiple application GUI elements that are spatially distant from each other on the screen (e.g., a comment and the associated document text). Therefore, the main design challenge for MagDocs was to find a way to seamlessly display the information in these GUI elements close to each other within the screen magnifier viewport, in order to reduce panning done by low vision users. The other design challenge was to facilitate convenient access to all the comments/changes in the document, given that these comments/changes are scattered all over the document and as such it was not easy to manually search for them via panning.

To address these challenges, MagDocs employs a system architecture shown in Figure 2. When a user invokes the MagDocs browser extension with a special shortcut, the `Control Manager` first takes a screenshot of the entire webpage and then passes it on to the `Segment Identifier`. The `Segment Identifier` then leverages custom extraction models to identify regions in the screenshot that correspond to comments or changes, and subsequently extracts text in the identified regions using optical character recognition (OCR). The motivation underlying the use of screenshot-based method was based on our observation that both the comments and changes have distinct visual patterns and screen locations that are clearly distinguishable from other regions of the Google Docs application. Next, the `DOM Subtree Identifier` uses the extracted text to detect webpage DOM subtrees that correspond to comments or changes. In parallel with the `DOM Subtree Identifier`, the `Text Location Detector` leverages the hierarchical content structure and the text highlighting aspect of Google Docs to identify the portions of the document text that have attached comments or changes. Next, the `Text Location Detector` maps the comments/changes with the identified text portions in the document based on their order of appearance in the application DOM tree. Once the text portions are mapped to the comments/changes, the MagDocs `Interface` finally creates a pop-up containing copied information from the first comment or change, and then displays this pop-up right above the corresponding text in the document as shown in Figure 1 (bottom row). The user can then either interact with the pop-up (e.g., post reply) or simply press a special shortcut to automatically move the focus to the text portion corresponding to the next comment/change. Furthermore, the `Control Manager` continuously monitors and instructs the `Text Location Detector` to update the locations of the pop-up interface in real time as necessary.

4.1 Control Manager

The Control Manager coordinates the operations of the MagDocs modules. It is specifically responsible for the following three tasks: *(i)* capturing the screenshot of a webpage and forwarding it to the Segment Identifier; *(ii)* monitoring the application to detect scroll events and accordingly intimating the Text Location Detector to update the position of the pop-up interface; and *(iii)* communicating with the browser to obtain the application DOM and then forwarding it to the DOM Subtree Identifier. The Selenium WebDriver [46] was used for capturing a webpage screenshot, and Flask API [42] was used to establish a communication channel between the MagDocs browser extension and Selenium submodule. The Chrome browser's native APIs were used to obtain the application DOM as well as to monitor the application for scroll events.

4.2 Segment Identifier

This module is responsible for two tasks: *(i)* identifying regions in the webpage screenshot that correspond to comments or changes; and *(ii)* extracting text from the identified regions.

The Segment Identifier leverages custom Mask R-CNN models [20] for identifying comments or revision regions in the application interface. Specifically, we use these models to identify regions of interest (ROIs) corresponding to the comments or revision details in the screenshot image of the webpage. This observation also holds for other similar collaborative writing applications such as Microsoft 365 Word Online. The Segment Identifier preprocesses the regions using Otsu's adaptive thresholding method [2] and exploits the Tesseract OCR service [17, 48] to extract all the text in these regions.

Dataset. To train Mask R-CNN models, we created two custom datasets for comments and changes, respectively, with each consisting of 450 document screenshots from collaborative writing tools along with ground-truth annotations. The two datasets were created by taking screenshots from Google Docs and Microsoft 365 Word Online collaborative writing tools (100 screenshots from each tool respectively). As manual annotation of large corpus is impractical, we applied image augmentation techniques [47] on the 200 screenshot images to artificially generate additional 250 images, resulting in 450 images for each dataset. The mask on each image that annotated the ROI corresponding to list of comments

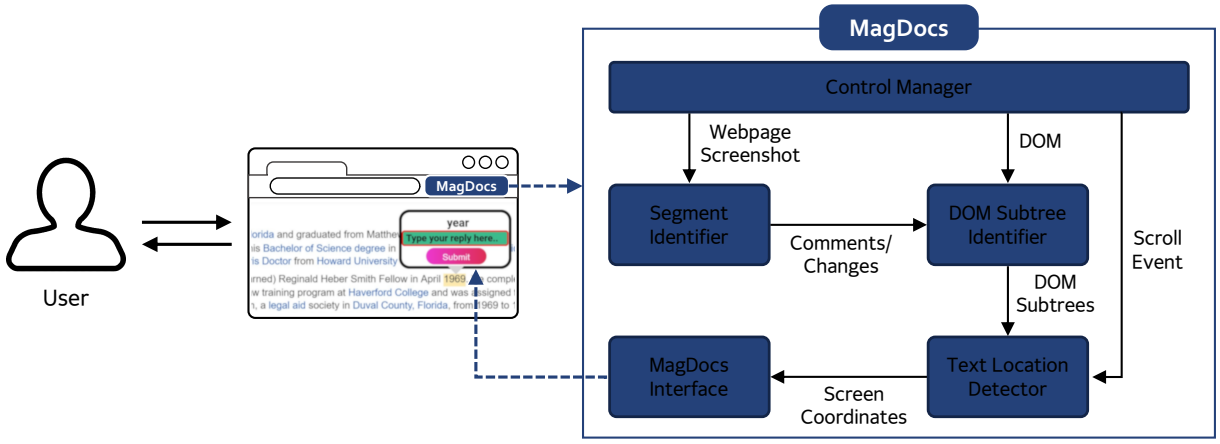


Figure 2: An architectural overview of MagDocs.

or changes was generated using the GIMP software[1]. From each dataset, we used 300 images for training, 130 for validation, and 20 for testing. All screenshot images were scaled to a standard size of 640×480.

Training environment and parameters. We trained each model for 4 epochs with 500 steps per epoch. The entire training process was accomplished with a single GPU (NVIDIA GeForce GTX 1650). Informed by prior works [20, 41], we trained the model in two phases as follows. In the first phase, we trained only head layers while freezing all the backbone layers with the learning rate of 0.001. In the second phase, we fine-tuned the entire model, i.e., trained all the layers, with the learning rate set to 0.0001. These parameter values were chosen to optimize the performance on validation set. We trained with both ResNet-101 and ResNet-50 as the backbone networks [21] along with Feature Pyramid Network [30].

Evaluation. We assessed the performance of our models using the standard Intersection over Union (IoU) and mean average precision (mAP) metrics. For the Mask R-CNN model capturing comments, its performance on the test set was found to be high – 91.2% (IoU) and 0.91% (mAP) with the ResNet-101 backbone network, and 83.6% (IoU) and 0.82% (mAP) with the ResNet-50 backbone network. Similarly, the performance of the model capturing revision details was also high – 89.1% (IoU) and 0.90% (mAP) with the ResNet-101 backbone network, and 84.5% (IoU) and 0.84% (mAP) with the ResNet-50 backbone network. As the ResNet-101 model yielded better results than ResNet-50 model, we integrated it into MagDocs.

4.3 DOM Subtree Identifier

The DOM Subtree Identifier determines the application DOM subtrees that correspond to the comments and revisions. First, the module tries to detect overlaps between the screen location information obtained from the Segment Identifier and the screen coordinates of individual DOM nodes (if available) in the DOM tree. If the information is not available, then it uses the comment/change text information obtained from the Segment Identifier to find the DOM nodes with matching inner text. The identified nodes are then passed on to the Text Location Detector.

4.4 Text Location Detector

This module is responsible for the following tasks: *(i)* determining text portions in the document that are associated with comments or revision; *(ii)* establishing a mapping between the identified text portions and the comments or changes detected by the DOM Subtree Identifier; and *(iii)* computing the location of the text on the screen to determine the appropriate position of the pop-up.

To determine the text portions associated with attached comments/revisions, the module exploits the DOM structure of Google Docs that hierarchically organizes the content of the document into pages, paragraphs, and lines. Given that Google Docs highlights portions of text which have attached comments (or changes in the 'Version History' application view), the Text Location Detector scans the immediate neighborhood of the individual 'line' nodes in the DOM tree to check if there are any special overlay nodes attached to them that highlight portions of those lines. If such overlay nodes exist, then the corresponding line nodes are guaranteed to contain comments/changes. Using the screen coordinate information available in these overlay nodes, the Text Location Detector is able to determine the exact text portions in the document where the comments are attached or changes have been made. The module also remembers the line numbers associated with these text portions so that appropriate adjustments can be made to the pop-up interface location, in case the user scrolls the document. To map the identified text portions with the corresponding comment/change DOM subtrees, the module leverages the relative positions of these nodes in the application DOM, i.e, the first comment subtree identified in the DOM (assuming Depth First Search traversal) is mapped to the first text portion identified in the DOM, and so on.

4.5 MagDocs Interface

This module is responsible for: *(i)* displaying the MagDocs pop-up interface at the appropriate location in the document; and *(ii)* interpreting user keystrokes and performing intended actions. The MagDocs interface supports the following special shortcuts:

[1]https://www.gimp.org/

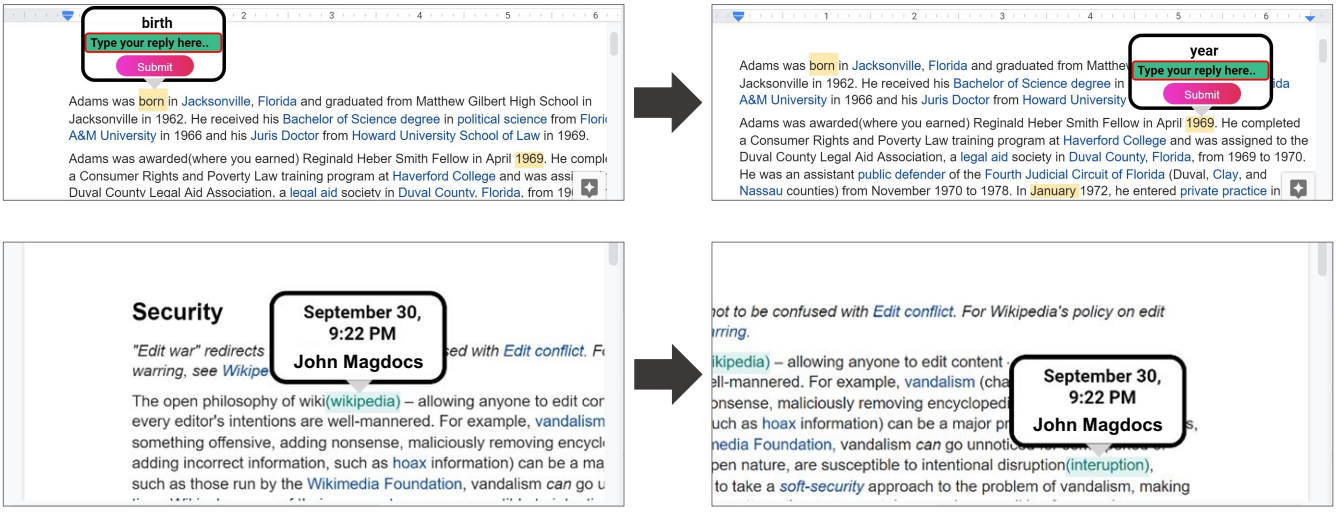


Figure 3: Illustration of MagDocs. *Top row*: navigating comments one-by-one; *bottom row*: navigating changes one-by-one.

- `CTRL` (Comments) – If a user is currently accessing a comment, shift focus to the next comment. If the pop-up is currently closed, activate the pop-up and move focus to the most recently viewed comment if it still exists, otherwise start from the first comment.
- `CTRL+N` (Changes) – If the user is currently checking a change in the version history, shift focus to the next change in the version history. If the pop-up is currently closed, activate the pop-up and move focus to the most recently viewed change.
- `SHIFT` – Close the pop-up (for both comments and changes).

An illustration of traversing comments and changes one-by-one via MagDocs is shown in Figure 3. By automating this access for screen magnifier users, MagDocs obviates the need for manual panning or scrolling to do the same – a major pain point for these users as uncovered in the interview study.

5 EVALUATION

5.1 Participants

To assess the MagDocs prototype, we recruited 15 low vision screen magnifier users through email lists and snowball sampling. The demographic information of participants is listed in Table 2 (Appendix B). The average age of participants was 47.9 (Median=49, Min=25, Max=65), and the gender distribution was almost the same (7 female, 8 male). As in case of the earlier interview study, we had the following inclusion criteria: *(a)* strictly using screen magnifiers – users with extremely poor visual acuity who rely on screen readers were excluded; *(b)* proficiency with the ZoomText [45] screen magnifier; and *(c)* familiarity with the Google Docs application. None of the participants used the color inversion feature or speech narration feature in the magnifier while doing the tasks. To ensure external validity, there was no overlap between the participant groups between this study and the interview study.

5.2 Apparatus

The user study was conducted using a Lenovo ThinkPad laptop running Windows 10 home edition, with the Google Chrome web browser and ZoomText screen magnifier installed. A traditional external keyboard and a mouse were plugged into the laptop. All participants indicated that they were familiar with the Windows 10 platform. The participants used experimenter-provided Google accounts which were created solely for the study.

5.3 Design

The participants were asked to do the two types of tasks on the Google Docs application: *(i)* **Task T1** – address all comments in a document by adding an appropriate Yes/No response; and *(ii)* **Task T2** – view all changes in an article and indicate the ones made by the collaborator 'John Magdocs' in the document text. In a within-subject experimental setup, the participants were asked to do the tasks under two conditions, with two tasks per condition: *(i)* **Screen Magnifier** (baseline) – the participants could only rely on the ZoomText screen magnifier; and *(ii)* **MagDocs** – the participants could leverage the MagDocs interface.

We created 4 Google Docs documents, 2 for each task. Specifically, for Task T1, each document was 1-page long and had 5 comments attached to arbitrary portions of the document text. However, between the two documents, the positions (i.e., line numbers) in the document where the comments were attached were roughly similar. The comment content was a question that the participants had to understand by referring to the associated text, and then respond to the comment by typing *Yes* or *No* in the text box below the comment. Similarly, for Task T2, the two documents were both 1-page long and had 6 prior edits, out of which 3 were performed by a dummy account with the name 'John Magdocs'. Specifically, the second, third, and fifth edits in both documents were made by 'John Magdocs'. Also, the positions (i.e., line numbers) of the edits in both documents were roughly similar. The pre-specified content for all the task documents was taken from the Wikipedia

articles about famous personalities in sports. The assignment of documents to tasks, and the ordering of tasks and conditions were counterbalanced with the Latin square method [6].

5.4 Procedure

The experimenter first allowed a participant to customize magnifier settings, such as zoom level and cursor enhancement. This was followed by a short practice session of about 20 minutes, where the participant was given enough time to get familiar with the MagDocs interface and its shortcuts. Then, the participant was asked to perform the study tasks in the predetermined order. The study concluded with a brief exit interview where the participant provided subjective feedback. With the participant's consent, the entire session was recorded. The experimenter also made notes regarding any peculiar interaction behavior or strategy exhibited by the participant while doing the tasks.

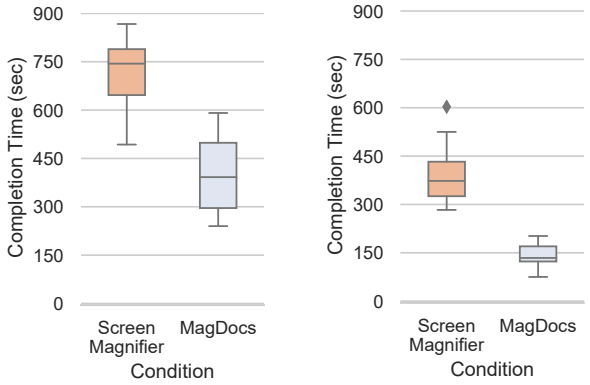
5.5 Data Collection and Analysis

We collected the following metrics and data: *(i)* task completion times; *(ii)* responses to the System Usability Scale (SUS) questionnaire [9] to evaluate perceived usability; *(iii)* responses to the NASA Task Load Index (NASA-TLX) questionnaire [19] for evaluating perceived user effort; and *(iv)* qualitative feedback from the participants as well as experimenter observations. As none of the tasks required the participants to edit the document content, the contribution of typing time towards the overall completion time was limited to the required responses (i.e., Yes/No) in Task T1. The subjective feedback from participants along with the experimenter's notes were analyzed using the open coding technique [43], where we iteratively scanned the data and identified recurring key insights and themes.

5.6 Results

Task completion times. Figure 4 presents task completion time statistics for the study tasks. In Task T1, the participants spent an average of 713.6 seconds (Median = 744, Min = 493, Max = 867) completing the tasks with the ZoomText screen magnifier, whereas they only needed an average of 412.7 seconds (Median = 392, Min = 240, Max = 591) with MagDocs. This difference in completion times was found to be statistically significant (Wilcoxon signed-rank test, $W = 1$, $z = -3.35$, $p = 0.0008$). All participants correctly answered all the questions in Task T1 under both study conditions. Similarly, for Task T2, the participants expended an average of 389.7 seconds (Median = 373, Min = 283, Max = 603) for successfully accomplishing the study tasks with ZoomText screen magnifier, whereas they took an average of 143.1 seconds (Median = 134, Min = 75, Max = 202) with MagDocs, which was significantly lower than that in the screen magnifier condition (Wilcoxon signed-rank test, $W = 0$, $z = -3.41$, $p = 0.0006$).

From these observations, it is clear that MagDocs significantly reduced the time required for addressing comments as well as reviewing changes for participants regardless of their visual conditions. A deeper analysis of the collected study data revealed that in the Screen Magnifier condition all participants spent a considerable amount of time panning to-and-fro between comments and the corresponding text portions in the document for Task T1 – this validates our earlier findings in the interview study. Moreover, for at least one comment in Task T1, all participants panned back-and-forth between the comments and the corresponding text portions more than two times. Also, the participants spent time locating the comments by manually panning and scrolling near the rightmost portion of the magnified screen content, and they also faced difficulty in linking comments with the corresponding text portions. In contrast, in the MagDocs condition, we noticed that the participants did not spend much time panning as they relied on the pop-up interface to view the comments right above the corresponding document text portions.



(a) Task T1 (b) Task T2

Figure 4: Completion time statistics for the study tasks.

We made similar observations regarding user interaction behavior for Task T2. In the MagDocs condition, all participants exploited the MagDocs pop-up interface to almost instantly determine the author of a particular change and also quickly navigate to the subsequent changes in the document. In fact, we noticed that much of the overhead in the MagDocs condition was due to the participants getting distracted by the changes, i.e., they spent some time reading what a change was about instead of focusing only on the name of the author as required by Task T2. However, in the Screen Magnifier condition, the participants spent considerable time panning between the document text and the changes (shown in the right pane as in case of comments). Compared to Task T1, the amount of back-and-forth panning by participants in the Screen Magnifier condition was significantly lesser in Task T2, presumably due to the fact that the participants only needed to determine the author name to complete the task.

Usability and perceived effort. We relied on the System Usability Scale (SUS) questionnaire [9] to evaluate usability of MagDocs and compare it with that of the status-quo screen magnifier. The SUS questionnaire consists of 10 alternating positive and negative 5-scale Likert statements, where a rating of 5 corresponds to *strongly agree* and 1 represents *strongly disagree*, with 3 representing a *neutral* rating. The responses to these 10 statements are then assimilated into a single score between 0 and 100, with higher scores indicating better usability. The average SUS score for the Screen Magnifier condition was $\mu = 58.33$ ($\sigma = 10.02$), which was significantly lower than that for the MagDocs condition

($\mu = 85.16$, $\sigma = 7.32$). Moreover, this difference in average scores was statistically significant (paired t-test, $t = 8.142$, $p < 0.001$).

To evaluate the perceived task workload, we used the NASA Task Load Index (NASA-TLX) questionnaire [19]. This questionnaire involves obtaining subjective rating feedback for six subscales: Mental Demand, Physical Demand, Temporal Demand, Overall Performance, Effort, and Frustration Level. The responses are combined into a single score between 0 and 100; however, unlike SUS, lower NASA-TLX scores indicate better performance. In our study, the average NASA-TLX score for the MagDocs condition ($\mu = 27.86$, $\sigma = 4.69$) was significantly lower than that for the Screen Magnifier condition ($\mu = 74.71$, $\sigma = 5.37$), with this difference being statistically significant (paired t-test, $t = -24.49$, $p < 0.001$). The analysis of the scores given to the individual subscales revealed that for the Screen Magnifier condition, the Mental Demand, Effort, and Frustration subscales (in that order) were the predominant drivers behind the high overall NASA-TLX scores. On the other hand, for the MagDocs condition, we observed that the distribution of scores across all the subscales was more uniform, and the scores were also much lower than those for the Screen Magnifier condition.

Qualitative feedback. The analysis of the subjective feedback from the exit interviews revealed following insights.

MagDocs is easy to learn and use. A majority (13) of participants attributed their high usability ratings to the simplicity and short learning curve of MagDocs. These participants expressed that remembering a few shortcuts to access and navigate the MagDocs pop-up interface was an *acceptable* trade-off, given the significant benefit of reduced panning effort.

Occlusion of document text is sometimes annoying. While all participants stated that the amount of frustration was significantly lower in the MagDocs condition than that in the Screen Magnifier condition, 6 participants indicated that they were slightly annoyed when the MagDocs pop-up interface occluded a relevant portion of the document text that they had to read in order to answer the questions in some of the comments. In this regard, 4 of these 6 participants further mentioned that they wished to be able to slightly move the pop-up interface based on their needs. Two participants also stated that they would like to be able to resize the pop-up interface as needed.

Customization of the MagDocs interface. Apart from resizing and repositioning the pop-up interface, the participants also suggested other customization options. For instance, 4 participants wanted to configure their own preferred keyboard shortcuts for accessing and navigating the pop-up interface. Three participants wished to alter text color, background, and border color for the pop-up interface. One participant even wanted to add more whitespace padding between the contents of the pop-up, even though it would possibly increase the panning effort.

6 DISCUSSION

6.1 Limitations and Future Work

A limitation of MagDocs is that it currently supports only Google Docs. This is because MagDocs relies on specific metadata in the DOM subtrees to determine the exact screen locations of document text portions that are relevant to these comments/changes, in order to properly display its pop-up interface close to them. Exploring mechanisms that can automatically learn where on screen to display the pop-up interface for comments/changes is the scope of our future work. Note however that the Mask R-CNN models for automatically detecting and extracting the comments/changes are generic, i.e., the models can be used for any arbitrary online collaborative writing applications other than Google Docs.

The present focus of MagDocs is also limited to comments and changes. While the interview study indicated that these two application features were most frequently accessed by most users for collaborative purposes, there are also other application features such as online chat that MagDocs might need to support in future for further enhancing the usability of online collaborative editors for screen magnifier users. MagDocs also does not currently support certain features that are exclusive to synchronous editing, e.g., mentioning active collaborators and their present cursor positions in the document.

Another limitation of MagDocs is that it is currently designed only for the Chrome web browser. Although Chrome is the most popular web browser among low vision users [51], there are still significant number of users relying on other browsers. Expanding MagDocs to support multiple browsers is mostly an engineering effort. Lastly, the sample sizes in both studies were small although they included a diverse set of low vision participants. Therefore, larger studies need to be conducted to validate our findings. Also, further work is required to understand the distinct needs and issues of individual low vision subgroups (e.g., peripheral vision, tunnel vision) so as to customize MagDocs for each subgroup.

6.2 Customization of MagDocs Interface

MagDocs currently offers a 'one-size-fits-all' interface with no support for customizing its interface. However, as reported in user study findings, many participants wanted to customize the MagDocs interface based on their individual preferences. These customizations included setting custom keyboard shortcuts for accessing and navigating the MagDocs pop-up interface, changing the relative position of the pop-up interface over the document text (e.g., *right* of the text instead of the default top of the text), resizing the pop-up interface, and so on. We plan to extend MagDocs to support these customizations.

7 CONCLUSION

Online collaborative writing applications have become increasingly prevalent in professional, academic, and even personal settings. We first investigated the interaction experiences of low vision screen magnifier users with these applications via an interview study, and found that the low vision users struggled to perform even basic collaborative tasks such as addressing collaborators' comments and reviewing document peer edits. As a first step towards addressing these issues, we proposed MagDocs, a browser extension for Google Docs, that enables low vision users to quickly and conveniently review document changes as well as address peer comments. A user study with low vision participants yielded promising results indicating significant improvements in both interaction efficiency and usability with MagDocs, when compared with the status-quo screen magnifier technology.

ACKNOWLEDGMENTS

We thank anonymous reviewers for their insightful feedback. This research work was funded by NSF awards 1805076, 1936027, 2113485, 2125147, and NIH awards R01EY030085, R01HD097188.

REFERENCES

[1] Aries Arditi and Jianwei Lu. 2008. Accessible Web Browser Interface Design for Users with Low Vision. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 52, 6 (2008), 576–580. https://doi.org/10.1177/154193120805200614 arXiv:https://doi.org/10.1177/154193120805200614

[2] Sunil L Bangare, Amruta Dubal, Pallavi S Bangare, and ST Patil. 2015. Reviewing Otsu's Method for Image Thresholding. *International Journal of Applied Engineering Research* 10, 9 (2015), 21777–21783.

[3] Jeffrey P. Bigham. 2014. Making the Web Easier to See with Opportunistic Accessibility Improvement. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. ACM, New York, NY, USA, 117–122. https://doi.org/10.1145/2642918.2647357

[4] Syed Masum Billah, Vikas Ashok, Donald E. Porter, and I.V. Ramakrishnan. 2018. SteeringWheel: A Locality-Preserving Magnification Interface for Low Vision Web Browsing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 20, 13 pages. https://doi.org/10.1145/3173574.3173594

[5] Yevgen Borodin, Jeffrey P. Bigham, Glenn Dausch, and I. V. Ramakrishnan. 2010. More than Meets the Eye: A Survey of Screen-Reader Browsing Strategies. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)* (Raleigh, North Carolina) *(W4A '10)*. Association for Computing Machinery, New York, NY, USA, Article 13, 10 pages. https://doi.org/10.1145/1805986.1806005

[6] James V. Bradley. 1958. Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design. *J. Amer. Statist. Assoc.* 53, 282 (1958), 525–528. https://doi.org/10.1080/01621459.1958.10501456 arXiv:https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1958.10501456

[7] Julian Brinkley and Nasseh Tabrizi. 2017. A Desktop Usability Evaluation of the Facebook Mobile Interface using the JAWS Screen Reader with Blind Users. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61, 1 (2017), 828–832. https://doi.org/10.1177/1541931213601699 arXiv:https://doi.org/10.1177/1541931213601699

[8] Cornelia Brodahl, Said Hadjerrouit, and Nils Kristian Hansen. 2011. Collaborative Writing with Web 2.0 Technologies: Education Students' Perceptions. *Journal of Information Technology Education* 10 (2011), IIP73 – IIP103. http://proxy.library.stonybrook.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=508425028&site=eds-live&scope=site

[9] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[10] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Giulio Mori. 2012. Designing E-Learning Collaborative Tools for Blind People. *E-Learning-Long-Distance and Lifelong Perspectives* (2012), 125–144. https://doi.org/10.5772/31377

[11] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, Giulio Mori, and Victor M. Penichet. 2014. Collaborative Editing: Collaboration, Awareness and Accessibility Issues for the Blind. In *Proceedings of the Confederated International Workshops on On the Move to Meaningful Internet Systems: OTM 2014 Workshops - Volume 8842*. Springer-Verlag, Berlin, Heidelberg, 567–573. https://doi.org/10.1007/978-3-662-45550-0_58

[12] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, Giulio Mori, and Victor M. R. Penichet. 2010. Accessing Google Docs via Screen Reader. In *Computers Helping People with Special Needs*. Springer Berlin Heidelberg, Berlin, Heidelberg, 92–99.

[13] Maitraye Das, Darren Gergle, and Anne Marie Piper. 2019. "It Doesn't Win You Friends": Understanding Accessibility in Collaborative Writing for People with Vision Impairments. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 191 (nov 2019), 26 pages. https://doi.org/10.1145/3359293

[14] Krzysztof Z. Gajos, Daniel S. Weld, and Jacob O. Wobbrock. 2010. Automatically generating personalized user interfaces with Supple. *Artificial Intelligence* 174, 12 (2010), 910–950. https://doi.org/10.1016/j.artint.2010.05.005

[15] Krzysztof Z. Gajos, Jacob O. Wobbrock, and Daniel S. Weld. 2007. Automatically Generating User Interfaces Adapted to Users' Motor and Vision Capabilities. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, Rhode Island, USA) *(UIST '07)*. ACM, New York, NY, USA, 231–240. https://doi.org/10.1145/1294211.1294253

[16] Google. 2022. Google Docs: Free Online Documents for Personal Use. https://www.google.com/docs/about/.

[17] Google. 2022. Tesseract OCR. https://github.com/tesseract-ocr/tesseract.

[18] Preston Gralla. 2020. G Suite vs. Office 365: What's the best office suite for business? https://www.computerworld.com/article/3515808/g-suite-vs-office-365-whats-the-best-office-suite-for-business.html.

[19] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139 – 183. https://doi.org/10.1016/S0166-4115(08)62386-9

[20] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2017. Mask R-CNN. In *Proceedings of the IEEE international conference on computer vision*. 2961–2969.

[21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. https://doi.org/10.1109/CVPR.2016.90

[22] Apple Inc. 2022. Change Zoom preferences for accessibility on Mac - Apple Support. https://support.apple.com/guide/mac-help/change-zoom-preferences-for-accessibility-mh40579/mac.

[23] Fawzi Fayez Ishtaiwa and Ibtehal Mahmoud Aburezeq. 2015. The impact of Google Docs on student collaboration: A UAE case study. *Learning, Culture and Social Interaction* 7 (2015), 85–96. https://doi.org/10.1016/j.lcsi.2015.07.004

[24] Julie A. Jacko, Armando B. Barreto, Gottlieb J. Marmet, Josey Y. M. Chu, Holly S. Bautsch, Ingrid U. Scott, and Robert H. Rosa, Jr. 2000. Low Vision: The Role of Visual Acuity in the Efficiency of Cursor Movement. In *Proceedings of the Fourth International ACM Conference on Assistive Technologies* (Arlington, Virginia, USA) *(Assets '00)*. ACM, New York, NY, USA, 1–8. https://doi.org/10.1145/354324.354327

[25] Richard L. Kline and Ephraim P. Glinert. 1995. Improving GUI Accessibility for People with Low Vision. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '95)*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 114–121. https://doi.org/10.1145/223904.223919

[26] Jonathan Lazar, Aaron Allen, Jason Kleinman, and Chris Malarkey. 2007. What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users. *International Journal of Human–Computer Interaction* 22, 3 (2007), 247–269. https://doi.org/10.1080/10447310709336964 arXiv:https://doi.org/10.1080/10447310709336964

[27] Cheuk Yin Phipson Lee, Zhuohao Zhang, Jaylin Herskovitz, JooYoung Seo, and Anhong Guo. 2021. CollabAlly: Accessible Collaboration Awareness in Document Editing. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 55, 4 pages. https://doi.org/10.1145/3441852.3476562

[28] Hae-Na Lee, Vikas Ashok, and IV Ramakrishnan. 2021. Bringing Things Closer: Enhancing Low-Vision Interaction Experience with Office Productivity Applications. *Proc. ACM Hum.-Comput. Interact.* 5, EICS, Article 197 (May 2021), 18 pages. https://doi.org/10.1145/3457144

[29] Lekha Limbu and Lina Markauskaite. 2015. How do learners experience joint writing: University students' conceptions of online collaborative writing tasks and environments. *Computers & Education* 82 (2015), 393–408. https://doi.org/10.1016/j.compedu.2014.11.024

[30] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. 2017. Feature Pyramid Networks for Object Detection. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 936–944. https://doi.org/10.1109/CVPR.2017.106

[31] Ahmad Zamri Mansor. 2012. Google Docs as a Collaborating Tool for Academicians. *Procedia - Social and Behavioral Sciences* 59 (2012), 411–419. https://doi.org/10.1016/j.sbspro.2012.09.295 Universiti Kebangsaan Malaysia Teaching and Learning Congress 2011, Volume I, December 17 – 20 2011, Pulau Pinang MALAYSIA.

[32] Natalie Maus, Dalton Rutledge, Sedeeq Al-Khazraji, Reynold Bailey, Cecilia Ovesdotter Alm, and Kristen Shinohara. 2020. Gaze-Guided Magnification for Individuals with Vision Impairments. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3382995

[33] Nihal Menzi-Çetin, Ecenaz Alemdağ, Hakan Tüzün, and Merve Yıldız. 2017. Evaluation of a University Website's Usability for Visually Impaired Students. *Univers. Access Inf. Soc.* 16, 1 (March 2017), 151–160. https://doi.org/10.1007/s10209-015-0430-3

[34] Microsoft. 2022. Free Microsoft Office Online. https://www.microsoft.com/en-us/microsoft-365/free-office-online-for-the-web.

[35] Microsoft. 2022. Use Magnifier to make things on the screen easier to see - Windows Help. https://support.microsoft.com/en-us/help/11542/windows-use-magnifier-to-make-things-easier-to-see.

[36] Lourdes Moreno, Xabier Valencia, J. Eduardo Pérez, and Myriam Arrue. 2018. Exploring the Web Navigation Strategies of People with Low Vision. In *Proceedings of the XIX International Conference on Human Computer Interaction* (Palma, Spain) *(Interacción 2018)*. Association for Computing Machinery, New York, NY, USA, Article 13, 8 pages. https://doi.org/10.1145/3233824.3233845

[37] Giulio Mori, Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Victor M. R. Penichet. 2011. Collaborative Editing for All: The Google Docs Example.

In *Universal Access in Human-Computer Interaction. Applications and Services*, Constantine Stephanidis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 165–174.

[38] Giulio Mori, Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Victor M. R. Penichet. 2011. Making "Google Docs" User Interface More Accessible for Blind People. In *Advances in New Technologies, Interactive Interfaces, and Communicability*. Springer Berlin Heidelberg, Berlin, Heidelberg, 20–29.

[39] Brian E. Perron and John Sellers. 2011. Book Review: A Review of the Collaborative and Sharing Aspects of Google Docs. *Research on Social Work Practice* 21, 4 (2011), 489–490. https://doi.org/10.1177/1049731510391676 arXiv:https://doi.org/10.1177/1049731510391676

[40] Shrinivas Pundlik, HuaQi Yi, Rui Liu, Eli Peli, and Gang Luo. 2017. Magnifying Smartphone Screen Using Google Glass for Low-Vision Users. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25, 1 (2017), 52–61. https://doi.org/10.1109/TNSRE.2016.2546062

[41] Hamed Raoofi and Ali Motamedi. 2020. Mask R-CNN Deep Learning-based Approach to Detect Construction Machinery on Jobsites. In *ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction*, Vol. 37. IAARC Publications, 1122–1127.

[42] Kunal Relan. 2019. Beginning with Flask. In *Building REST APIs with Flask*. Springer, 1–26.

[43] Johnny Saldaña. 2021. *The coding manual for qualitative researchers.* sage.

[44] John Gerard Schoeberlein and Yuanqiong Wang. 2014. Usability Evaluation of an Accessible Collaborative Writing Prototype for Blind Users. *J. Usability Studies* 10, 1 (nov 2014), 26–45.

[45] Freedom Scientific. 2022. ZoomText Screen Magnifier and Screen Reader - zoomtext.com. https://www.zoomtext.com/.

[46] Selenium. 2022. Selenium. https://www.selenium.dev/.

[47] Connor Shorten and Taghi M Khoshgoftaar. 2019. A Survey on Image Data Augmentation for Deep Learning. *Journal of big data* 6, 1 (2019), 1–48.

[48] Ray Smith. 2007. An Overview of the Tesseract OCR Engine. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Vol. 2. 629–633. https://doi.org/10.1109/ICDAR.2007.4376991

[49] Sarit Felicia Anais Szpiro, Shafeka Hashash, Yuhang Zhao, and Shiri Azenkot. 2016. How People with Low Vision Access Computing Devices: Understanding Challenges and Opportunities. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility* (Reno, Nevada, USA) *(ASSETS '16)*. ACM, New York, NY, USA, 171–180. https://doi.org/10.1145/2982142.2982168

[50] Mirza Muhammad Waqar, Muhammad Aslam, and Muhammad Farhan. 2019. An Intelligent and Interactive Interface to Support Symmetrical Collaborative Educational Writing among Visually Impaired and Sighted Users. *Symmetry* 11, 2 (2019), 238.

[51] WebAIM. 2018. Survey of Users with Low Vision #2 Results - WebAIM. https://webaim.org/projects/lowvisionsurvey2/.

A PARTICIPANT DEMOGRAPHICS FOR THE INTERVIEW STUDY

ID	Age/ Gender	Diagnosis	Visual Acuity		Screen Magnifier	Collaborative Writing Tools
			Left Eye	Right Eye		
P1	34/F	Leber congenital amaurosis	20/400	20/400	ZoomText	Docs
P2	36/M	Retinitis pigmentosa	20/500	0	ZoomText, Apple Zoom	Docs, Sheets
P3	51/M	Optic atrophy	20/400	20/400	ZoomText, Windows Magnifier	Docs, Microsoft 365 Word
P4	60/M	Glaucoma	0	20/500	ZoomText	Docs
P5	49/F	Chorioretinal scarring	20/400	20/400	ZoomText	Docs, Sheets
P6	31/F	Stevens-Johnson syndrome	20/200	20/200	ZoomText, Apple Zoom	Docs, Microsoft 365 Word
P7	62/M	Glaucoma	20/400	20/400	ZoomText	Docs
P8	56/M	Diabetic retinopathy	20/400	20/400	ZoomText	Docs
P9	32/F	Chorioretinal scarring	20/400	20/200	ZoomText, Apple Zoom	Docs
P10	43/M	Albinism	20/100	20/100	ZoomText, Apple Zoom	Docs, Sheets

Table 1: Participant demographics for the interview study.

B PARTICIPANT DEMOGRAPHICS FOR THE USER STUDY

ID	Age/ Gender	Diagnosis	Visual Acuity		Screen Magnifier	Collaborative Writing Tools
			Left Eye	Right Eye		
P1	65/M	Glaucoma	20/500	0	ZoomText	Docs
P2	42/M	Stargardt macular degeneration	20/200	20/200	ZoomText, Windows Magnifier	Docs, Microsoft 365 Word
P3	49/M	Albinism	20/100	20/100	ZoomText, Apple Zoom	Docs, Sheets
P4	46/M	Optic atrophy	20/200	20/200	ZoomText	Docs, Sheets
P5	62/F	Glaucoma	20/400	0	ZoomText	Docs
P6	35/F	Optic atrophy	20/400	20/400	ZoomText, Apple Zoom	Docs, Sheets
P7	52/M	Diabetic retinopathy	20/400	20/400	ZoomText	Docs
P8	59/F	Glaucoma	0	20/400	ZoomText	Docs
P9	40/F	Stargardt macular degeneration	20/200	20/200	ZoomText, Windows Magnifier	Docs, Microsoft 365 Word
P10	63/F	Glaucoma	20/400	0	ZoomText	Docs
P11	51/M	Retinitis pigmentosa	20/400	20/400	ZoomText	Docs
P12	33/F	Stevens-Johnson syndrome	20/200	20/200	ZoomText, Apple Zoom	Docs, Sheets
P13	25/M	Albinism	20/100	20/100	ZoomText, Apple Zoom	Docs, Sheets
P14	59/M	Glaucoma	20/400	20/400	ZoomText, Apple Zoom	Docs
P15	38/F	Cone dystrophy	20/400	20/400	ZoomText, Apple Zoom	Docs, Sheets

Table 2: Participant demographics for the user study evaluating MagDocs.

Visual-Meta Appendix

The data below is what we call Visual-Meta. It is an approach to add information about a document to the document itself, on the same level of the content (in style of BibTeX). It is very important to make clear that Visual-Meta is an approach more than a specific format and that it is based on wrappers. Anyone can make a custom wrapper for custom metadata and append it by specifying what it contains: for example @dublin-core or @rdfs.
The way we have encoded this data, and which we recommend you do for your own documents, is as follows:
When listing the names of the authors, they should be in the format 'last name', a comma, followed by 'first name' then 'middle name' whilst delimiting discrete authors with ('and') between author names, like this: Shakespeare, William and Engelbart, Douglas C.
Dates should be ISO 8601 compliant.
Every citable document will have an ID which we call 'vm-id'. It starts with the date and time the document's metadata/Visual-Meta was 'created' (in UTC), then max first 10 characters of document title.
To parse the Visual-Meta, reader software looks for Visual-Meta in the PDF by scanning the document from the end, for the tag @{visual-meta-end}. If this is found, the software then looks for @{visual-meta-start} and uses the data found between these tags. This was written September 2021. More information is available from https://visual-meta.info for as long as we can maintain the domain.

@{visual-meta-start}

@{visual-meta-header-start}
@visual-meta{version = {1.1},
generator = {ACM Hypertext 21},
organisation = {Association for Computing Machinery}, }

@{visual-meta-header-end}

@{visual-meta-bibtex-self-citation-start}

@inproceedings{10.1145/3511095.3531274,
author = {Lee, Hae-Na and Prakash, Yash and Sunkara, Mohan and Ramakrishnan IV and Ashok, Vikas},
title = {Enabling Convenient Online CollaborativeWriting for Low Vision Screen Magnifier Users},
year = {2022},
isbn = {978-1-4503-9233-4},
publisher = {Association for Computing Machinery},
address = {New York, NY, USA},
url = {https://doi.org/10.1145/3511095.3531274},
doi = {10.1145/3511095.3531274},
abstract = {Online collaborative editors have become increasingly prevalent in both professional and academic settings. However, little is known about how usable these editors are for low vision screen magnifier users, as existing research works have predominantly focused on blind screen reader users. An interview study revealed that it is arduous and frustrating for screen magnifier users to perform even the basic collaborative writing activities, such as addressing collaborators' comments and reviewing document changes. Specific interaction challenges underlying these issues included excessive panning, content occlusion, large empty space patches, and frequent loss of context. To address these challenges, we developed *MagDocs*, a browser extension that assists screen magnifier users in conveniently performing collaborative writing activities on the Google Docs web application. MagDocs is rooted in two ideas: *(i)* a custom *support interface* that users can instantly access on demand and interact with collaborative interface elements, such as comments or collaborator edits, *within* the current magnifier viewport; and *(ii) visual relationship preservation*, where collaborative elements and the corresponding text in the document are shown close to each other within the magnifier viewport to minimize context loss and panning effort. A study with 15 low vision users showed that MagDocs significantly improved the overall user satisfaction and interaction experience, while also substantially reduced the time and effort to perform typical collaborative writing tasks.},
numpages = {11},
keywords = {Online Collaborative Writing, Low Vision, Screen Magnifier, Visual Impairment, Accessibility, Assistive Technology},
location = {Barcelona, Spain},
series = {HT '22},
vm-id = {10.1145/3511095.3531274} }

@{visual-meta-bibtex-self-citation-end}

@{visual-meta-end