

CollabAlly: Accessible Collaboration Awareness in Document Editing

Cheuk Yin Phipson Lee*
Carnegie Mellon University
Pittsburgh, PA, USA
cheukyil@andrew.cmu.edu

Zhuohao Zhang*
University of Washington
Seattle, WA, USA
zhuohao@uw.edu

Jaylin Herskovitz
University of Michigan
Ann Arbor, MI, USA
jayhersk@umich.edu

JooYoung Seo
University of Illinois
Urbana-Champaign
Champaign, IL, USA
jseo1005@illinois.edu

Anhong Guo
University of Michigan
Ann Arbor, MI, USA
anhong@umich.edu

ABSTRACT

Collaborative document editing tools are widely used in professional and academic workplaces. While these tools provide basic accessibility support, it is challenging for blind users to gain collaboration awareness that sighted people can easily obtain using visual cues (e.g., who is editing where and what). Through a series of co-design sessions with a blind coauthor, we identified the current practices and challenges in collaborative editing, and iteratively designed CollabAlly, a system that makes collaboration awareness in document editing accessible to blind users. CollabAlly extracts collaborator, comment, and text-change information and their context from a document and presents them in a dialog box to provide easy access and navigation. CollabAlly uses earcons to communicate background events unobtrusively, voice fonts to differentiate collaborators, and spatial audio to convey the location of document activity. In a study with 11 blind participants, we demonstrate that CollabAlly provides improved access to collaboration awareness by centralizing scattered information, sonifying visual information, and simplifying complex operations.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Accessibility technologies**.

KEYWORDS

Collaborative writing, collaboration awareness, accessibility, Blind, visual impairment, screen reader, voice font, spatial audio, earcon

ACM Reference Format:

Cheuk Yin Phipson Lee, Zhuohao Zhang, Jaylin Herskovitz, JooYoung Seo, and Anhong Guo. 2022. CollabAlly: Accessible Collaboration Awareness in Document Editing. In *CHI Conference on Human Factors in Computing*

*Both authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution-Share Alike International 4.0 License.

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9157-3/22/04.
<https://doi.org/10.1145/3491102.3517635>

Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3491102.3517635>

1 INTRODUCTION

Collaborative document editing tools such as Google Docs, Microsoft Word, and Overleaf have become ubiquitous in today's professional and academic workplaces. Using shared documents, collaborators can synchronously or asynchronously access and modify document content, manage different versions, and track feedback and comments. Prior work has studied collaborative writing practices and observed a variety of editing strategies, including mixing synchronous and asynchronous writing to increase efficiency [9, 41]. This is made possible by features of document editors that provide collaboration awareness, such as shared cursors, editing histories, and comments. The collaboration awareness afforded by these tools allows users to decide what parts of a document to work on and when based on role structures and privacy needs [52].

However, collaborative document editing is challenging for blind users as these tools are often not sufficiently accessible [2, 12, 45]. Although commercial tools such as Google Docs and Microsoft Word have screen reader support for collaboration information in various text interface panels, navigating to this information is difficult, and in many cases, users are not made aware that there is new information available. This information also tends to lack the document context that the visual versions provide, increasing the mental load of understanding collaborator actions [12]. While sighted users can take advantage of the visual markup alongside the document text (i.e., highlights, nested comments, color-coded cursors) to quickly and passively obtain a rich awareness of the activities and changes, these benefits are not replicated non-visually [12]. For example, while Google Docs automatically provides verbal announcements when a user enters a line with a comment on it, it does not mention who made the comment and when; Google Docs also provides collaborator announcements when someone enters or leaves the document, but it is difficult for screen reader users to know what they are looking at or editing at a given time. Obtaining this information disrupts the document editing process and makes it challenging for blind users to have up-to-date collaboration awareness, if at all. Thus, existing accessibility support lacks the continuous and contextualized awareness of collaborator actions that visual cues provide.

In this work, we aim to address this issue through the design and development of CollabAlly, a system that makes collaborator actions and context accessible for blind users. To design CollabAlly, we first investigated how blind users currently accomplish collaborative writing tasks in Google Docs. Through a series of co-design sessions with a blind coauthor who has years of collaborative editing experience, we find that potentially useful collaboration information is usually scattered in the complex user interface of Google Docs. This, combined with the difficulty to remember hot keys, focus switching issues, and unclear information mean that blind users hardly fully make use of these collaborative functions. Through this process, we identified a set of challenges and requirements that informed the design of CollabAlly.

Based on these findings, we iterated on the design of CollabAlly using a prototype system as a means to explore various auditory representations of collaborative cues in practice. Das et al. identified that audio features such as earcons and multiple text-to-speech (TTS) voices are useful in understanding collaborative cues [13], and that spatial audio could be used for further disambiguation [12]. Thus, we used these representations as a basis for our initial prototype. We aimed to investigate and refine these representations within the context of a fully responsive extension to a popular collaborative writing tool. Through our co-design sessions, we iterated on the presentation and method for accessing these features, eventually generating a design for our full system, CollabAlly.

CollabAlly is a system consisting of both a front-end browser extension and a backend server to extract and maintain information about a collaborative document. By identifying key class structures and HTML elements in the document, CollabAlly parses and tracks collaborators' behaviors, comment updates, and text changes. CollabAlly presents this information both in real-time using earcons for background events, and on-demand using a browser-injected dialog box. To further contextualize who, where, and when each of these events have occurred in the document, CollabAlly employs spatial audio and voice fonts, thus giving blind users a comprehensive understanding of the collaborative environment.

Through a user study with 11 blind participants who are regular users of Google Docs and collaborative writing, we demonstrate that CollabAlly is effective at providing access to collaborative editing features of Google Docs. Using CollabAlly, many participants were able to envision future changes to their document editing workflows that would benefit them in their personal and professional lives. CollabAlly demonstrates the potential in using a variety of audio representations to support collaboration awareness. From our evaluation, we also highlight future usability improvements and considerations for accessible collaboration systems.

2 RELATED WORK

Our work is related to prior work on (i) collaborative writing tools and practices, and (ii) collaborative editing accessibility. In designing our system, we also take inspiration from (iii) prior systems for accessible collaboration in other areas, and audio-based virtual exploration techniques for screen reader users. CollabAlly is intended to apply audio-based features to make collaborative document editing more accessible for screen reader users.

2.1 Collaborative Writing

Researchers have long studied collaborative writing practices, including document production strategies and group feedback dynamics [5, 6, 9, 29, 42]. Early research in this area explored techniques such as shared edits, color-coded users, and collaborator locations in collaborative writing systems (for example, ShrEdit [42], Quilt [17], and SASSE [3]) that have influenced commercial writing applications today, such as Microsoft Word [34] and Google Docs [23]. A portion of this research has also focused on document production and editing strategies used among certain groups, and the group dynamics that influence this writing behavior [5, 28, 29, 39, 50, 51]. Boellstorf et al. describe how collaborators alternate between synchronous and asynchronous editing to balance contributions and increase writing efficiency [9]. Similarly, Olson and colleagues have studied students' collaborative writing practices in a variety of scenarios and observed a range of outlining and hand-off strategies [43, 57]. Wang et al. observed that people are hesitant to edit content directly written by their supervisors, using document history to avoid doing so [52]. Birnholtz and Ibara found that people added comments in shared documents explaining their actions to others to avoid misinterpretation [6].

A large portion of the collaborative writing systems, frameworks, and behaviors that have been studied over the years are enabled by features of collaborative document editors intended to support *collaboration awareness*. Dourish and Bellotti present collaboration awareness as the shared understanding of others' actions and edits that is necessary to coordinate and provide a context for individual efforts [14]. The theory of social translucence suggested by Erickson and Kellogg proposes awareness as a key component to group functioning, as making social information visible leads to a group accountability for that information in their actions [16]. Thus, collaborative writing systems often attempt to facilitate awareness by providing collaborators equal access to information [3, 17, 26, 42].

A variety of experimental tools over the years have presented such features, and popular collaborative editing tools Microsoft Word and Google Docs both also provide awareness features for facilitating collaboration. For example, Google Docs provides real-time shared cursor locations that are color-coded so editors can easily understand collaborator actions, shared comments that appear next to the text that is commented on for context, and a full color-coded document history for awareness of detailed changes [23]. While these features are incredibly useful and enable a range of collaborative editing behaviors that support efficiency and group dynamics, they are often communicated solely through visual markup: colors, highlights, floating comment locations, and strikethroughs. In this work, we aim to develop non-visual alternatives to provide blind users with a level of collaboration awareness that was previously inaccessible.

2.2 Collaborative Writing Accessibility

While commercial collaborative writing tools Microsoft Word and Google Docs have basic screen reader support [20, 35, 36], this support is often not enough to facilitate collaboration awareness, or even provide easy access to collaborative editing features [11, 47, 49]. Prior work has studied the experiences of screen reader users working with collaborative writing software. Das et al. present an

interview study of blind collaborators in ability-diverse teams, and identified a variety of barriers that they face when trying to access collaboration features [12]. These include high learning curve and no learning resources, and difficulty accessing different commenting menus. Information about collaborator actions was also found to be cognitively overloading because it lacked context; for example, a screen reader announcing a comment might simply say ‘comment entered’, or it might read out collaborators’ changes character by character. Screen reader users then have to do the additional work of understanding who is doing what and where.

Some prior work has begun to develop prototypes for improving the accessibility of collaborative editing features, such as tracking document changes. For example, Schoeberlein and Wang developed a prototype of an accessible revision interface as a Word add-in, which consisted of a pop-up window which would read the paragraph context, then the text change [46, 47]. Waqar et al. created a collaborative writing system which used speech input and audio alerts for text modifications [53, 54]. These prototypes highlight the importance of research to make collaboration awareness accessible. We aim to build upon this work by creating a fully working system for accessible collaboration that can be used directly in Google Docs. We take inspiration from these prior designs, and aim to extend upon them by providing additional context for each collaborator action, and an easier navigation mechanism.

Most relevant to our work is a set of auditory representations for collaboration activity designed by Das et al. [13]. They use a variety of audio techniques, including non-speech audio and different voice presentations, to convey comments and recent edits in a document to screen reader users. Their work demonstrates that audio representations are promising in signifying collaborative writing activities, and provides suggestions for which techniques could be used in different contexts. In CollabAlly, we take inspiration from this work and integrate similar signals into an end-to-end system that augments Google Docs. We further create a full system architecture for extracting collaborative information from documents and presenting audio signals in context. This allows us to further study such signals within the typical document editing context.

2.3 Accessible Collaboration and Navigation

In this work, we also take inspiration from prior accessible system design. A variety of systems have been developed to support collaboration awareness for blind and visually impaired people in other contexts, for example programming [2], games [19, 55, 56], and education [32, 38]. Audio feedback is a common design element of such systems. For example, Mendes et al. designed an accessible interactive tabletop system that successfully used spatialized speech to convey target locations and voice fonts to distinguish collaborators [33]. Prior work has similarly used spatial audio in various navigation applications, for example, Microsoft Soundscape [37] among others [1, 4, 8], or in gaming applications [56]. CollabAlly similarly implements a variety of audio feedback, in this case, within the document editing context, for identification and localization of collaborator actions.

3 CO-DESIGN PROCESS

In this section, we discuss the stages of our co-design process. The primary goals of our co-design process were to create a shared

understanding of the current practices and challenges that screen reader users have in collaborative editing, and to inform system iteration in the next stage. First, we conducted a series of co-design workshop sessions with a blind coauthor where we sought to understand current challenges faced when using Google Docs for collaborative writing and verify our design requirements for our system. Second, together we iterated on the design of an initial prototype system aiming to improve Google Docs accessibility. We identified issues with our current design ideas and brainstormed new ones via discussion. The result of our co-design process was a set of design considerations. We present the co-design process, the design considerations that we drew from the process, and explain how they informed our final system, CollabAlly.

3.1 Co-design Method

We worked with our blind coauthor (referred to as JY) to understand the detailed challenges faced in collaborative editing, current workflows, and how we may design a system to address these challenges that fits into existing workflows.

3.1.1 Co-Designer Background. JY, a coauthor on this paper, collaborated with the remaining authors to develop a shared understanding of the specific challenges and barriers to writing collaboratively with Google Docs. JY is blind, while the remaining authors are sighted. JY is an academic who has approximately 6 years of experience using Google Docs for academic writing and collaborative editing. He collaborates with others (both blind and sighted collaborators) on Google Docs very often for notes and manuscripts, with a frequency of at least 2-3 times a week. He also has other collaborative editing experience, using tools like Microsoft Word for writing and Visual Studio Code for pair programming. JY uses a screen reader (primarily on the Windows operating system with NVDA and JAWS), and also uses a refreshable Braille display.

3.1.2 Apparatus. As a basis for our co-design sessions, we used both Google Docs as-is, and a modified Google Docs environment we created to address some of the issues JY and prior work had identified. This initial prototype was iterated on after the first and second section to incorporate the findings of those sessions.

First, we used Google Docs with its existing accessibility support. We asked JY to turn on screen reader support and collaborator announcements in Google Docs accessibility settings, which are the two primary accessibility features built into Google Docs.

Second, we created an initial prototype system that served as an extension to collaborative writing in Google Docs. Prior to the co-design sessions, we incorporated findings from a study by Das et al. that aimed to understand accessibility in collaborative writing [12] to create this initial prototype, including the recognized complexities in collaboration awareness. For example, sighted people can use visual cues like comments, track changes, real-time editing notifications in Google Docs while screen reader users need to hear all of notifications alongside the document text. To address these complexities, we considered providing more contextual information to blind users. We drew the concept of spatial audio from augmented and virtual reality environments where the document is considered as a flat plane and notifications come from different

spatial directions. Based on the idea, we designed an initial prototype that used spatial audio and text-to-speech (TTS) to convey various activity within Google Docs. The prototype pulled information from the document on who was currently collaborating on the document, what comments were left by whom and where, and what recent edits had been made. We will discuss the details of the prototype and how we iterated on it during the co-design process in Section 3.3. This prototype was implemented as a Chrome extension, and further details on our final complete system, CollabAlly, are included in Section 4.

3.1.3 Procedure and Analysis. We had three formal co-design sessions over three weeks with JY, followed by numerous consultations in weekly meetings on our system's features as we continued to iterate. The first session lasted 2.5 hours, while the other two lasted 1.5 hours.

In the first session, we focused on understanding the current practices and challenges JY had when using Google Docs. We asked JY to use Google Docs with its accessibility settings to accomplish a series of collaborative writing tasks, and then installed our initial prototype system to do similar tasks in another document. These tasks focused on making sense of collaboration activities in the document, including (i) locating collaborators' positions in text and their usernames, (ii) understanding comments left by collaborators, and (iii) understanding text changes made by collaborators. We use the term 'understanding' here in a broad sense, and note that it can include understanding who performed the activity, what the change was, where the change was in the document, and the surrounding document context. For example, in one task, two researchers joined the document and left comments at separate locations. We then asked JY to find the comments, find the highlighted comment text, and also to find the paragraph and document section that the comments were in.

In the next two sessions, we focused on testing and iterating on the features of our initial prototype system. We began these sessions with a short introduction of the session's purpose and a description of our initial prototype's current functionality and how it had progressed since the last session. We spent time discussing each feature of Google Docs related to collaboration (e.g., navigating to collaborators, accessing comments, tracking recent changes) and our initial prototype to create a shared understanding and brainstorm ideas. These discussions in our co-design sessions were open and loosely structured to make sure JY could express anything he thought of. In between these sessions, members of the research team implemented JY's suggestions so that they could be tested in the following session.

All authors were present in these co-design sessions. The sessions were video recorded, and the authors alternated taking notes throughout all sessions. Two researchers then met online to discuss the notes. We extracted key insights that reflected JY's challenges and strategies, which we present in the following sections.

3.2 Current Practices and Challenges

We first discuss the current practices and challenges reported by JY throughout the three co-design sessions. Note that although we mainly learned the reported practices and challenges in the first session, the other two sessions also resulted in new insights on this

topic as JY used our prototype. We put particular emphasis on what information blind people need for better collaborative editing, and how they want to access that information from an expert's (JY) perspective. Later in the full user study of our system, we confirmed these challenges with our other participants.

3.2.1 Identifying and Locating Collaborators. Knowing what collaborators are present in a shared document is an important task for JY. As highlighted by prior work, people writing collaboratively often change their behaviors based on social and organizational roles of their collaborators [12], thus, this is a key task in collaborative editing. However, JY reported that there was not an efficient way to accomplish the task of identifying and locating collaborators in Google Docs. When JY was asked to locate collaborators using Google Docs' accessibility settings, he tried to navigate to the 'live edit' panel that displays edit history after he entered the document. However, he then realized that only when collaborators made edits will he be able to find out about their information. JY then reported that *"I have no idea what or where people are in the doc."* JY also mentioned that he knew all about Google Docs' collaborator announcement feature where it *"just announces when a collaborator's cursor is at the same spot as me, or when a collaborator leaves my line,"* and when he wanted to actively find out about the current collaborators, he lost track of them and said that *"I have no idea how I would handle this."* Furthermore, since the announcements happen involuntarily, JY was more inclined to skip through them by pressing the screen reader shortcuts.

One workaround mentioned by JY was to scan the entire document line-by-line to hear collaborator announcements. While this is possible, it is tedious, especially for long documents. Another workaround to access this information is to navigate to the collaborators profile pictures at the top right corner of the page, which would read out collaborator names. However, JY noted that these are difficult to reach, as they require the user to navigate out of the document text area. In other words, even though it was technically possible for him to get to the information through complex navigation, JY preferred not to do that because it was not well-integrated into his existing workflow.

JY emphasized that his overall goal is to know who is in the document on-demand, instead of the status quo where blind people get rounds of collaborator announcements for everyone 'coming and leaving' which is distracting when he is performing other tasks. He added that it is distracting because it requires a significantly greater cognitive load to perceive and interpret what is being said. This issue was also identified by Das et al. [12] where the serialized presentation of text-based content and collaborative awareness information were cognitively overloading. This suggested that our solution should provide information about what collaborators are in the document and at what location in an easy-to-navigate and on-demand fashion.

3.2.2 Tracking Text Changes. Tracking text changes in a shared document is one of the most important tasks for collaborative editing, since it is the most fundamental feature of collaborative editing in a common space. When we asked JY to locate text changes which were just made in the document, he reported that *"this will be very tricky,"* and that *"Google Docs by nature does not provide a way for screen reader users to identify recent changes."* JY said he knows that

“there is a version history option, but it is hard to use and I never tried to really use it in my work.” Apart from version history, a workaround blind users have is to open the ‘live edit’ panel. However, as was mentioned in identifying collaborators (Section 3.2.1), navigating outside of the document is not the preferred method for JY to incorporate into his existing workflows. Another workaround is to use suggestion mode to track changes. However, the changes will still be scattered throughout the document, which require line-by-line reading. JY mentioned that he used to export them to native Word documents so that he could see all of the changes there in one accessible place, however it adds additional workload to his workflow. Similar behavior of pulling up a list of changes or comments had been identified by Das et al. in their studies [12]. Similar to identifying and locating collaborators, tracking text changes is technically possible for blind people, but it usually requires complex operations and was not preferred in blind users’ workflows. This suggested that our solution should provide an integrated summary of text changes and can be expanded to detail exploration and navigation through the changes in documents.

3.2.3 Working with Comments. Interacting with comments from different collaborators is also an integral part of collaborative editing. Despite being a commonly desired feature, JY reported that interacting with comments in Google Docs is not very accessible. When using Google Docs with accessibility settings to find comments, JY was experienced and knew the keyboard shortcuts for locating and replying to comments. He also demonstrated the way to use the ‘help’ option in Google Docs’ menu bar and search the keyword ‘comment,’ which listed all the comment-related operations including moving to next or last comment. However, JY could not easily identify the corresponding text that a comment was on; his typical workaround for this was to go over the document paragraph by paragraph and reply to comments he encountered. JY also mentioned that for blind users who are not proficient at remembering all the keyboard shortcuts, they would have to use the help function and search for available features, which is tedious. And even though they could eventually complete the task, the process was not easy to use and can be significantly improved. This suggested that our solution should provide information that gives blind users an awareness of the comment’s textual context, and a method for them to easily navigate to the comment or its surrounding text.

3.2.4 Making use of Contextual Information. Similarly to Das et al. [13], we refer to contextual information as both the textual context of an action (i.e., the surrounding content of a comment or text change), and the collaborative context (i.e., a collaborator’s current position and past actions). Besides co-designing our prototype through a series of task-based usability discussions, we also considered other design choices such as how to make use of contextual information in shared documents. JY expressed that it was important to access the location where collaboration activities occurred rather than simply knowing what activity occurred. Additionally, to make full use of contextual information, JY wanted to have a solution deeply integrated with various modalities beyond screen readers only, including using non-speech audio, TTS, and Braille displays. This suggested that our solution should consider contextual information access and multi-modal presentation of such information.

3.3 System Iteration

Prior to the first session with JY, we implemented an initial prototype system for collaborative writing accessibility that presents visual information about collaborators, comments, and text changes using multiple audio representations. We iterated this prototype with JY over the co-design sessions and we present the process of iteration here.

3.3.1 On-demand Access to Information. Our initial prototype used automatic speech announcements when certain events happened. For example, when a collaborator left a comment, our prototype used TTS to announce who left a comment and at what location. When using this feature, JY asked that if *“there is a command to stop reading the audio?”* Furthermore, a major concern JY had is that he *“cannot review the information that is spoken on the fly,”* and he had to *“memorize this information, and listen to it all again to see what I missed.”* When a text change was announced, JY mentioned, *“I remember that there was a color change and a font size change, but I don’t remember what the text change was because I had to memorize lengthy verbal information.”* Our initial prototype also used a set of four keyboard commands to provide on demand access to information about collaborators, comments, and recent text changes. When pressed, an entire summary of collaborative activities would be read to the user via TTS. However, these keyboard commands were difficult to remember, and the long audio summaries suffered from the same issues as the automatic updates.

After rounds of discussion on this matter, JY and the researchers designed a dialog box that can co-exist with any existing screen reader. The dialog box integrates all desired information related to collaboration activities and can be accessed on demand. It also saves the extra cognitive load of remembering interactive keyboard commands by using the native screen reader to read through information. To use the dialog box, blind users only need to remember one keyboard shortcut to open it, and can then read the information in the dialog box using standard screen reader navigation. Not only is this more familiar, but it also provides the ability for users to skip over information that is not relevant by jumping to different headings in the text. Another reason we chose to use the dialog box is that by presenting information in text rather than speech, it can be accessed using other modalities such as Braille displays, which has the potential to be used by deaf blind users as well.

When designing the dialog box, we also decided to keep some information as background notifications. This is driven by prior work on providing collaboration awareness, and also through our co-design sessions where we concluded that interrupting normal screen reader flow is detrimental. Thus, in a dialog box we provide access to collaborator, comment, and text change details. In the background, we also provide earcons for collaborators joining and leaving the document, new comments, and for a collaborator enters or leaves the paragraph that the user is editing. While the first two are simplified versions of features that we provide in the dialog box, we chose not to provide a background notification for text changes as we believed it would be too distracting. Inspired by Das et al. [12], who highlight the need for understanding when you are writing ‘on top of’ someone else’s work and understanding when you are being watched, we chose to notify users when someone was in close proximity to their cursor.

3.3.2 Spatial Audio Iteration. One of the major features we implemented in our initial prototype was to use spatial audio to relay the relative location of collaborators' locations and activities. The document was treated as a two-dimensional space and the visual flow of collaboration activities was mapped to audio directions. For example, if the blind user is at the very beginning of the document, and a collaborator makes a comment at the bottom right corner of the first page, the blind user would hear announcements coming from their bottom right direction. This was designed to give blind users more spatial awareness of where the collaboration activity happened so that they can react to it properly. However, when JY was prompted to respond to an activity that was farther away, he mentioned that he could recognize that it was a lower volume so it was farther away, but he could not find the position from that. He also raised concerns about using volume to indicate distance. He mentioned that *"the 3D audio is enough (for mapping to a specific location in a document). Don't change the volume since we rely on audio information. It will be hard if we can barely hear it."* Furthermore, he noticed that the spatial directions of top and bottom were harder to distinguish than the directions of left and right. Instead, the shared document can be viewed as a linear structure since the horizontal positions in a single line is less important than vertical positions in the whole document. Therefore, JY and the researchers decided to save two directions (left and right) only to indicate spatial positions. For example, an announcement coming from the right direction indicates that it is located below the blind user in the document. Blind users can also rely on how left or right the announcements are to tell how far the activities are in the document. However, the volume stays the same for the matter of clarity.

3.3.3 Audio Voice Fonts. In the initial version of our prototype, we used different voice fonts to distinguish between different collaborators' actions. For example, a comment might be read out with a gendered voice with an American English accent. However, when we changed to use the dialog box, we decided to make this feature optional. JY described how this could conflict with normal screen reader usage, and that layering an additional audio source without the user's permission could cut things off or cause the loss of information. In the dialog box, users can still access a version of the text read with voice fonts by pressing a button.

3.3.4 Trade off Between TTS and Non-speech Audio. In our prototype, we also used a mix of non-speech audio and TTS techniques for audio representation. For example, we assigned non-speech audio for each collaboration activity and played text announcements afterwards. However, JY pointed out that TTS should be considered optional because of the potential conflicts with the default TTS in screen readers. For current screen readers like JAWS, the announcements can be cut off with keyboard commands. JY mentioned a couple of times when he wanted to turn off the TTS audio or the native screen reader announcements and focus on the other audio source that he was interested in. This implies that we should consider the trade-off between TTS and non-speech audio. On the one hand, non-speech audio is implicit and it requires training and memorization before using, but it is also less distracting and interrupting when the native screen reader is playing. On the other hand, TTS provides accurate announcements and blind users need such information about the collaboration activities or their contexts.

Based on the consideration, JY and researchers talked about the potential of reducing cognitive loads using implicit audio, since blind users would not need to focus on other verbal TTS announcements, and if they find the implicit audio represents something worth looking, they can use TTS audio as another option.

Motivated from our co-design process and the study of audio representations in collaborative writing by Das et al. [13], we iterated our prototype to integrate non-speech and TTS audio. We used non-speech audio as the main source for collaboration activity notification while we positioned the access of TTS audio in the dialog box where users can click for more information. For example, when a collaborator left a comment, instead of directly reporting TTS audio, an implicit audio will be played and the blind user can open up the dialog box to look into the corresponding tab. The details of the user interaction can be found in Section 4.

3.4 Design Considerations

Extracting from the insights generated from our co-design process, we propose six design considerations that we use to guide a system design that makes collaboration awareness accessible.

Support common collaboration activities. Our solution should provide information about collaborators, comments, and text changes. Also, a summary of these changes are preferred. This overall consideration is both driven by prior research, and powered by our co-design process with JY. How JY struggled to accomplish the writing tasks in the study sessions also proved that it was necessary to support these collaboration activities.

Provide both on-demand and automatic updates. Our solution should communicate collaborative information through both an on-demand method, and an automatic, real-time method. We created a dialog box for blind users to access information on-demand. For real-time information, we design background notifications that would not conflict with screen reader audio.

Provide contextual information. Our solution should provide the appropriate amount of information for users to make sense of collaboration activities. We observed that with Google Docs, information about where comments were left in the document was difficult to access, and providing just the highlighted text was not enough, as comments often only highlight a single word.

Provide easy navigation. Our solution should provide navigation from collaboration activities to their corresponding places in text to avoid (i) document traverse and (ii) switching between the document area and other areas like comment or edit history panel. We observed that with Google Docs, it was difficult to jump from reading a comment to making changes in the text in response to it.

Support various modalities. Our solution should be deeply integrated with existing tools like non-speech audio, TTS, and Braille displays, etc. This is motivated by both prior research [13] and our co-design process.

Simplify operations. Our solution should make complex keyboard shortcuts more simplified and usable. JY reported in the sessions that too many keyboard commands would be harder to remember for blind users.

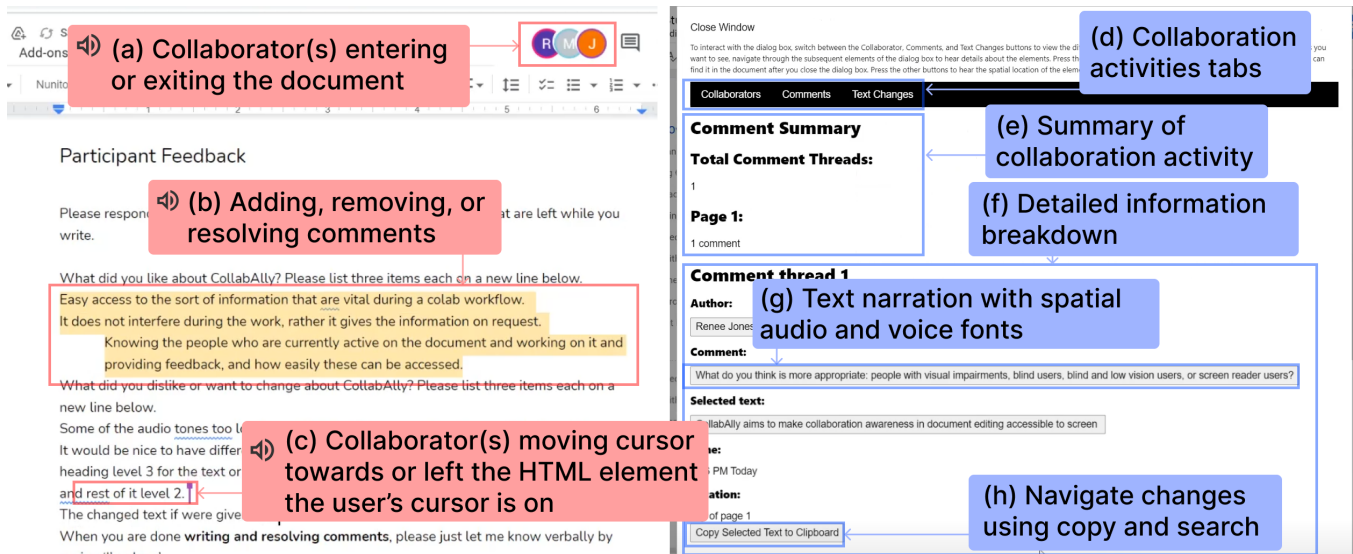


Figure 1: The main user interactions in CollabAlly, comprising of earcons that will play automatically based on real-time collaboration activities in the document (left); and an on-demand dialog box interface that parses and formats all collaboration activities for blind users to navigate using existing screen reader commands (right).

4 COLLABALLY SYSTEM

Drawing from our iterative co-design process, we designed CollabAlly, a system that aims to enhance collaboration awareness and make collaborative document editing more accessible for blind users [30]. CollabAlly comprises a front-end browser extension that tracks activities in the collaborative environment and conveys it to users through a combination of earcons [7] and a dialog box; and a backend server that stores collaboration activities in the document. In particular, CollabAlly also aims to provide more contextual information about the collaborative environment, including who, where and when each collaboration activity has occurred, through a combination of audio features including non-speech audio, spatial audio, and voice fonts. In this section, we first describe how users might interact with CollabAlly as they edit documents collaboratively, and then detail the framework under which the system was implemented. CollabAlly is open sourced at: <https://github.com/HumanAILab/CollabAlly>

4.1 User Interaction

Figure 1 illustrates the user interactions of CollabAlly. Results from our co-design process indicated that our solution should provide access to contextual information and support different modalities (see *Design Considerations* in Section 3.4). To that end, we designed a set of earcons that concisely conveyed real-time updates about collaborators and comments as users are editing. Furthermore, to integrate existing collaboration activities and provide regular on-demand updates, we designed a dialog box interface that users can navigate through a combination of predefined keyboard shortcuts, as well as screen reader keyboard commands blind users access on a regular basis. The dialog box works by popping up in browser and offers comprehensive information about all collaboration activities

and organizes them in an intuitive, hierarchical format. Furthermore, leveraging the capabilities of earcons and spatial audio, users can receive real-time feedback about collaborator activities and comment updates in the document without being disrupted from their ongoing document editing. Finally, users also have access to collaborator-defined voice fonts to read activities in the voice that each collaborator has selected, thereby enhancing their awareness of each collaborator's behavior and distinguishing different changes in the document.

Note that there are several technical challenges that affected our implementation because of how documents are structured in Google Docs: (i) document and real-time collaborator information is only available by accessing the HTML DOM tree and parsing desired elements, as the Google Docs API is limited to basic reads and writes, (ii) only visible content on the browser screen can be fetched in real time, (iii) the complex user interface and nested DOM structure makes it difficult to navigate between HTML elements or change users' cursor focus.

4.1.1 Implicit Audio. Findings from our co-design process indicated that users highly valued regular updates about collaborators and comments in the document as they are working, both to ensure they do not interfere with collaborators' work, but also to increase their awareness of what collaborators are doing in the document (see Section 3.2.1). Blind users did not like existing collaborator announcements that use TTS, as they were distracting. However, they also wanted access to real-time updates about the document (see Section 3.4). As illustrated in Figure 1a, b, and c, with CollabAlly, users have access to a set of six distinct earcons, which represent whether a collaborator has (i) joined or (ii) left the document; whether a comment was (iii) added or (iv) deleted; and whether the collaborator's cursor has (v) moved into, or (vi) away from the user's current editing element. As an example, if a

user turns on CollabAlly and collaborators have entered the document, a harp earcon sound will play. Similarly, if a user is editing a paragraph and a collaborator moves their cursor to the paragraph, the user hears a distinct beep sound informing them that a collaborator's cursor is nearby, either to read or to make modifications to the paragraph. Due to technical challenges (i) and (ii) from Section 4.1, CollabAlly treats certain events as identical, such as (i) a collaborator who is idle and a collaborator who has left; and (ii) a resolved comment and a deleted comment. Hence, these events map to the same earcon and are not distinguished by our current implementation.

To further contextualize where the collaborator is positioned in the document, CollabAlly also uses spatial audio to pan the earcon either to the left or right of the user, indicating that the collaborator has moved their cursor above or below the user, respectively. The same spatial mapping is used for all earcons, and users have the option of customizing the directionality of this spatial mapping in CollabAlly's settings menu. With earcons, these updates are communicated automatically and succinctly, without requiring blind users to actively search for them, thereby are less disruptive to their existing document editing tasks.

4.1.2 Change Summary. Beyond what earcons can communicate, blind users also wanted on-demand access to information about different collaboration activities. For example, after hearing the earcon indicating that a collaborator has left the document, users may want to check how many collaborators are still active. With CollabAlly, users can quickly access this information by opening the CollabAlly dialog box, which contains information about collaborators, comments, and text changes. We found that users typically wanted to first access a summary of the activities in the document, without having to spend excessive time reading and scrolling through to identify the information scattered around (see Section 3.4). To address this need, once a user accesses the tab for a specific type of activity they want to examine, they can read through a summary of those changes at the top of the dialog box to get an overview, as illustrated in Figure 1d and e. For instance, users can navigate to the 'Collaborators' tab in the dialog box and identify the number of collaborators that are active, who moved, left, or are idle under the 'Total Collaborators' subheading. Similar rules apply to the 'Comments' and 'Text Changes' tabs as well. Summaries, along with all other elements in the dialog box, are organized by subheadings that adhere to W3C's accessibility standards [15].

4.1.3 Change-Specific Contextual Information. In addition to the summaries for each of the three collaboration activity categories, as users scroll down the dialog box, they can access further details about each type of change, as illustrated in Figure 1f. For example, as a collaborator's cursor moves close to the user, the user may want to know who's cursor it is, and what line they are reading. To access this information, users return to the dialog box, access the 'Collaborators' tab, and navigate down the collaborator details below the summary section. Under the details section, users have access to each collaborator's name, the location of their cursor in the document, as well as the line of text that their cursor is hovering over. Using this information, users can develop a mental model of where each collaborator is located, and infer which collaborator is near their cursor and what they are looking at.

Similarly, users may want to read through the comment threads in the document, to keep track of the discussions and conversations among other collaborators, or to see whether a collaborator has responded to a comment that the user has made. By accessing the comment details below the 'Comments' summary, users can read through the author of each comment; the content of the comment itself; the highlighted text corresponding to each comment; the time that each comment was posted; the location of the highlighted text; and all corresponding replies in chronological order.

Finally, to track text changes in the document, users can navigate below the summary under the 'Text Changes' tab, to access each text change in chronological order. Each entry in the dialog box describes what the change was; the corresponding text that was changed; the location of the change; what the text was before the change; and what the text is after the change. In addition, any visual, stylistic changes to a piece of text is also listed under the 'Text Changes' tab. The details for a style change follow an identical format as other text changes, but instead, describes the visual elements of the text before and after the change was applied. For example, if a collaborator changed the font size and color of a sentence, users would access both the font size and color of the sentence before and after the collaborator applied the stylistic changes. This method provides an alternative to version control, where users previously had to access older versions of the document and manually compare the text to identify what changes were made (see Section 3.2.2). In addition, CollabAlly provides blind users with access to information pertaining visual and stylistic changes that were previously unavailable in Google Docs. This provides blind users with further context about what each collaborator is working on, and further paints a mental model of the changes in the document.

All of this information further enhances users' collaboration awareness in the document. Furthermore, because this navigation method leverages blind users' existing screen reader commands, it requires minimal additional learning to use and access. As we learned in the co-design process, users expressed hesitation using Google Docs and their built-in accessibility features because they require extensive additional keyboard shortcuts that they have to memorize (see Section 3.2.3). CollabAlly remedies this by only using keyboard shortcuts for opening the dialog box and settings menu.

4.1.4 Spatial Audio and Voice Fonts for Contextual Details. To distinguish the changes based on the collaborator it corresponds to, CollabAlly embeds spatial audio and voice fonts to further contextualize information about comments and collaborators in the dialog box. As blind users access the selected text for each collaborator, they can press the button containing the selected text (as illustrated in Figure 1g), which will play a narration of the selected text in a distinct voice representing the collaborator, spatially positioned based on whether the text is above or below the user's cursor. Similarly, users can also activate the buttons under comment and text change details to have them read out using spatial audio and the voice of the original author. We selected voice fonts from Google Cloud Text-to-Speech [25]. For prototyping purposes, voice fonts were pre-selected for the specific collaborators, but in practice, we envision that collaborators can choose their own voice fonts in advance through their local version of the CollabAlly browser extension, to best represent themselves in a collaborative setting.

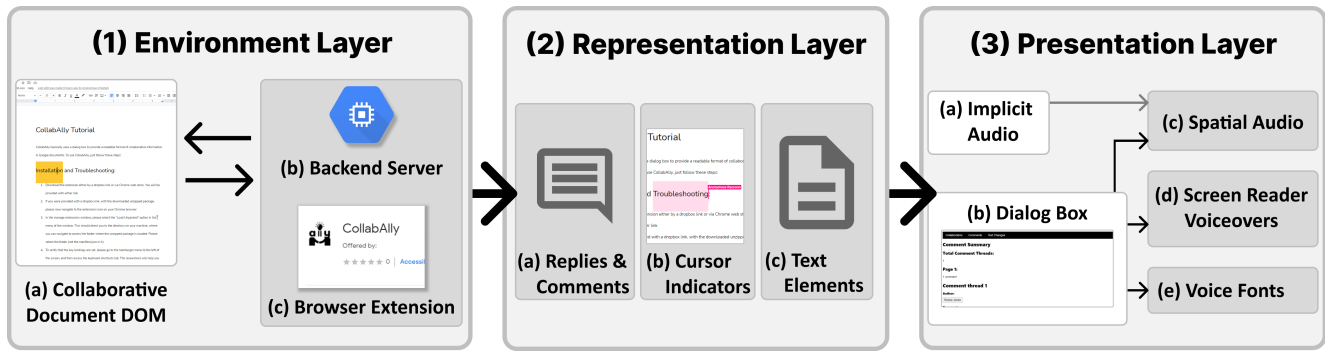


Figure 2: CollabAlly’s three-layer model system architecture, comprising (1) the Environment Layer that extracts the collaborative environment, e.g., the Google Docs DOM; (2) the Representation Layer that fetches and parses collaboration activities; and (3) the Presentation Layer that presents the contextual information to blind users, e.g., using audio features.

4.1.5 Navigation through Copy and Search. Beyond reading and identifying where certain activities are located, blind users may also navigate their cursors to them in the document. To do so, users can press the ‘Copy to Clipboard’ button to copy the selected text, and navigate to that element using the built-in ‘search text’ feature, as illustrated in Figure 1h. Due to the technical challenges (ii) and (iii) mentioned in Section 4.1, CollabAlly is unable to automatically reposition and refocus users’ cursors to the selected text, thus leading to this workaround. Nonetheless, with better support for accessing HTML elements and manipulating cursors in Google Docs in the future, navigation can be further streamlined.

4.1.6 Asynchronous Activity Updates. CollabAlly supports asynchronous activity updates after users leave and rejoin a document. When a user returns and activates CollabAlly, they will have access to the text changes and updated comments in the dialog box since they last accessed the document. CollabAlly achieves this by storing all collaboration activities in a persistent database.

4.2 System Architecture

Figure 2 illustrates the system architecture with which CollabAlly was implemented. As our goal is to make CollabAlly potentially extensible across different use cases beyond Google Docs and document editing in the future, our system architecture is application-agnostic and comprises of three layers, each of which is used to extract and model specific types of information necessary to capture collaborator behavior and states in real time, and then communicate that information to blind users in an accessible manner and format.

4.2.1 Environment Layer: Extracting the Collaborative Environment.

To establish a representation of the collaborative environment in which the user is interacting with, we established a foundational layer for CollabAlly that fetches an instance of the collaborative application itself, as illustrated in Figure 2. For Google Docs, this was done through the browser extension (Figure 2.1c), by fetching the DOM tree in real time using Javascript (Figure 2.1a). In addition, to track changes in the collaborative environment, CollabAlly used a backend server that fetched a MobileBasic version of the document (Figure 2.1b), which was used to compare different versions of the document to parse any text changes. Through this representation of the collaborative environment, CollabAlly can then proceed to

identify and parse key collaboration activities and state information that blind users want as they interact collaboratively in the application.

4.2.2 Representation Layer: Fetching and Parsing Collaboration Activity.

The next layer involves querying the document itself and parsing visual elements from the document into the models from the Environment Layer. We did so by reverse engineering the structures and changes of a document from its DOM tree in real time. Through our co-design process, we identified the specific information blind users want and need to know as they edit documents, and mapped their needs into data structures containing specific information about collaborators, comments, and text changes, as described in Section 4.1. To ensure that users could collaborate both synchronously and asynchronously, we also defined a backend system and database used to store this information persistently when users exit the document. The server and the database were implemented using Python and MongoDB, respectively. Upon re-entering the document and opening CollabAlly, the system can query the remote server and database to access and later parse the asynchronous changes.

To extract the necessary information to communicate to blind users, we identified different class names of desired HTML elements (i.e., collaborator cursors and comments) to locate them in the document, and added a loading screen to scroll over the entire document so that non-visual elements can also be fetched (related to technical challenge (ii) in Section 4.1). In particular, we referenced HTML elements used to visualize (i) user replies and comments (Figure 2.2a); (ii) collaborators’ cursors (Figure 2.2b); and (iii) text elements in the document (Figure 2.2c). We store and track the different states of collaborators, comments, and document text, and use this data to provide both real-time support, and asynchronous activity updates.

CollabAlly intermittently queries the shared document to identify the cursor positions of collaborators, comment boxes with their corresponding highlighted texts, and text changes. Then, CollabAlly maintains an updated state of these elements and compares it with previously queried states. Specifically, CollabAlly uses Google’s ‘Diff-Match-Patch’ library [21] to analyze text changes. The resulting changes are then reformatted using a library we developed that

parses style and content changes, to be communicated to users in a digestible format. Each change is mapped to the page number and location in the document, to help users locate and navigate to it.

4.2.3 Presentation Layer: Presenting Contextual Information Using Audio Features. The Presentation Layer maps the collaboration activities extracted by the Representation Layer into specific presentations, such as audio features. CollabAlly uses spatial audio and earcons to support document navigation by representing text locations as audio locations relative to the user’s cursor location (Figure 2.3a and c). For on-demand information, CollabAlly injects a dialog box into the document (Figure 2.3b), which also supports spatial audio while also allowing end users to navigate through detailed information using built-in screen reader announcements (Figure 2.3d). All audio repositioning is implemented using the Resonance Audio SDK [24].

Furthermore, CollabAlly supports voice fonts (Figure 2.3e). Each collaborator is assigned a unique voice ID that can be pronounced through different TTS synthesizers. Currently, this feature is natively available for screen readers supporting Nuance Vocalizer engine [40]. For those using non-vocalizer-supported screen readers, CollabAlly also provides voice fonts through Google Cloud’s TTS interface to generate server-side speech.

5 USER STUDY

The goal of our user study was to evaluate CollabAlly with a larger portion of the blind community to understand how it supports collaborative writing and awareness after system iteration, and how it fits into or complements existing user workflows. Overall, our method used for the user study is task-based usability testing. We sought to understand the strengths and limitations of CollabAlly for various tasks that involves collaboration activities.

5.1 Participants and Apparatus

We recruited 11 blind participants from an email list for blind academics (see table 1). Participants were between 19 and 58 years of age ($\mu = 34.8$ years, $\sigma = 13.2$ years); 6 were male, 4 were female, and 1 was trans/non-binary. Among them, 8 participants were fully blind, and 3 had low vision and were unable to read any print. All participants used screen readers in order to access their devices, and all had some prior experience using Google Docs. Participants reported using Google Docs for a variety of purposes including in their role as students (4 participants), as part of their job (6 participants), and for their personal writing (1 participant).

In our study, P1 and P2 used Google Docs and P3-P11 used CollabAlly as the apparatus. Apart from the apparatus, all participants were given all the same instructions, including being introduced to the study, being asked to rate how important collaborative activities were to them, and being instructed to complete a same series of collaborative writing tasks, with a post-study interview at the end. We report the results of P1 and P2 using Google Docs and P1-P11’s responses to pre-study interviews in Section 6.1 to present the current challenges in collaborative writing using Google Docs. We then report results from the remaining 9 participants (P3-P11), who used CollabAlly as the apparatus, in the remaining subsections.

Participants installed CollabAlly as a Chrome extension, and shared their screen over Zoom while completing the study tasks.

Participants used their typical desktop screen reader (6 used NVDA, 4 used JAWS, and 1 used VoiceOver) during the study with their chosen speed and language settings.

During the studies using CollabAlly, we asked participants to turn on Google Docs screen reader support, but turn off Google Docs collaborator announcements as these conflicted with CollabAlly’s. These settings were reverted and the extension was removed at the end of the study. Our study was approved by our institution’s IRB, and participants consented to participating in the study through both email and verbal consent at the beginning of the study session.

5.2 Procedure

Participants were first asked a series of demographic and background questions. Participants were then asked to rank their agreement on a seven-point scale with statements on the importance of four common tasks in collaborative writing: knowing what collaborators are at the document, locating collaborators’ cursors, working with comments, and tracking recent changes. For each of these statements, we asked participants to explain their rating, and to explain their current strategy for completing the task in Google Docs, if they had one. Our aim with these questions was to understand how collaboration awareness is important for blind users.

Participants then installed the CollabAlly Chrome extension, and were given a walkthrough of the system on a practice document. Participants read through each panel in CollabAlly’s pop up window, were prompted to try CollabAlly’s spatial audio buttons, and were able to ask any questions about CollabAlly’s functionality. The study administrator also played CollabAlly’s six background notification earcons for participants twice and explained the meaning of each sound.

Participants were then asked to complete a series of tasks as described below in Section 5.3. For each task, one of the study administrators opened the study document with two study Google accounts with made-up names (Renee Jones and Marcos Valdez). The study administrator then made comments and changes for each task under these accounts. Participants could ask questions, give feedback, or explain their actions at any point during the tasks. After each task, participants were asked to rate their agreement with two statements on a seven-point scale, including “This task was easy for me to complete” and “The information provided by the system was helpful for completing this task.” Participants were asked to explain their ratings for each statement, and were then asked to give open-ended feedback about the practicality of various features if used in their typical writing workflow. Each session took approximately 2 hours to complete, and participants were compensated \$25 per hour for their participation.

5.3 Tasks

We designed the following tasks to mimic common collaborative writing tasks. All of the tasks were completed in the same order for each participant. Each of the tasks was performed on the same document content.

Task 1: Locating Collaborators. Participants were asked to locate two collaborators in the document, and describe their locations

Table 1: Participants' demographics, including their apparatus used, gender, age, self-identified visual ability, operating system and screen reader used in the study.

Participant ID	Study Apparatus	Gender	Age	Self-identified Visual Ability	Operating System of Use	Screen Reader of Use
P1	Google Docs	Female	58	Blind	Windows	JAWS
P2	Google Docs	Male	43	Blind	Windows	NVDA
P3	CollabAlly	Male	19	Blind	Windows	NVDA
P4	CollabAlly	Female	31	Blind	Windows	JAWS
P5	CollabAlly	Male	58	Blind	Windows	JAWS
P6	CollabAlly	Male	26	Blind	Windows	NVDA
P7	CollabAlly	Female	27	Low Vision	Windows	NVDA
P8	CollabAlly	Male	30	Blind	Windows	NVDA
P9	CollabAlly	Male	27	Blind	Windows	NVDA
P10	CollabAlly	Trans/Non-binary	25	Low Vision	Mac OS	VoiceOver
P11	CollabAlly	Female	39	Low Vision	Windows	JAWS

with as much detail as possible (above or below current location, page number, line of text, surrounding text, section heading).

Task 2: Identifying Text Changes. Participants were asked to locate two text changes in the document (edits were made directly in the document, not with suggestion mode), describe the change with as much detail as possible (author, type, content, etc.), and describe the locations of each change with as much detail as possible.

Task 3: Working with Comments. Participants were asked to locate two comments in the document, describe them with as much detail as possible (author, type, content, etc.), and describe the location of each comment with as much detail as possible. The first comment was made on a single line of text, and participants were asked to reply to it. The second comment was made on an entire paragraph, overlapping the first comment, and participants were asked to make changes to the paragraph according to the comment.

Task 4: Collaborative Writing. Participants were asked to write six sentences in response to two prompts at the bottom of the document (“List three things you liked about CollabAlly” and “List three things you disliked about CollabAlly”). As they wrote, study administrators moved around the document, and added four comments to trigger CollabAlly’s background sound notifications. Two of the comments were on a few words of text in a single line, typically asking for clarification or explanation, and two of the comments were over three lines, asking participants to add a number next to each line in order to indicate the importance of each item to them. Participants were instructed to use the comments to edit their responses as they saw fit.

5.4 Analysis

Since participants used the study time to fully explore the functions, ask usability questions, and give feedback, performing a statistical analysis of task completion time does not provide much insight about how CollabAlly works in practice. Instead, we report qualitative data on participants’ strategies and workflows during task completion and collaborative writing using CollabAlly compared to situations without CollabAlly, and their general feedback.

We first transcribed the study sessions from the screen recordings, additionally logged what information was found by participants for each task, and created written descriptions of participants’ task completion strategies. We then noted the completion rate of all tasks, including the number of participants who were able to identify various types of information, for example, the number of

participants who could identify the page number or section heading associated with a change.

For our qualitative analysis, two members of the research team analyzed the study sessions using thematic analysis as described by Braun and Clarke [10]. Participants’ study transcripts, along with the written descriptions of their task completion strategies, were treated as data items to identify trends in participant feedback. We first individually read and familiarized ourselves with the data. We performed an open coding of the data independently, then adjusted the codes as a group until sufficient agreement was reached. We focused on identifying themes relating to participants’ task completion strategies, and commentary on their existing editing workflows.

6 RESULTS

In this section, we describe our study results including how participants used CollabAlly to complete each task and common feedback.

6.1 Current Google Docs Challenges

Participants were asked to describe the importance of three common tasks in collaborative writing. Here, we present the results of those descriptions, and common challenges faced by participants when attempting to complete these tasks. We also present their responses to Likert-scale questions of rating importance of collaborative tasks in Figure 3. We use this to confirm our earlier findings from our co-design sessions, and to underscore the importance of providing accessible collaboration awareness in Google Docs.

6.1.1 Identifying and Locating Collaborators. 9 out of 11 participants believed that knowing which collaborators are in the document is important to them when writing collaboratively, and 8 out of 11 participants believed that locating collaborators’ exact cursor locations is important when writing collaboratively. Participants used this information to coordinate and avoid interfering with others’ work. For instance, P9 mentioned that because different colleagues have different editing styles, “*knowing who is collaborating on the document can clue me in as to what I might be looking for later on, if I’m going to be looking for notes or formatting changes.*” P6 described avoiding collaborators edits: “*If I know that person is working, I can back off.*”

However, all participants reported that they lacked an efficient way to accomplish this task. They typically used Google Docs

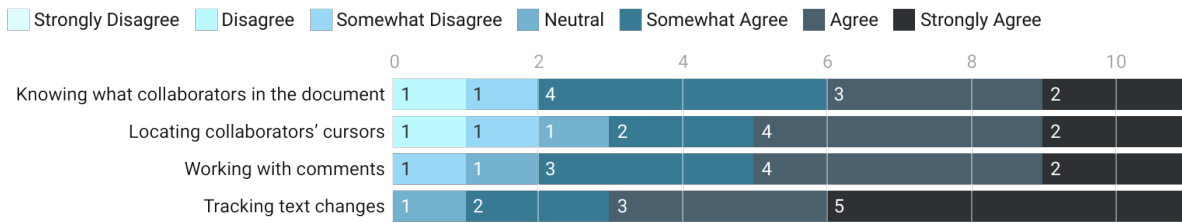


Figure 3: Participants' responses to how important different collaborative writing tasks were to them, indicating that most participants felt important to know collaborator, comment, and text change information.

collaborator announcements, which announce when a collaborator joins or leaves the document or the paragraph they are editing, though these require additional attention and are hard to keep track of. One common work around we observed was participants (P3, P5, P8) using Google Docs' line-by-line collaborator announcements to scan the entire document for collaborators. While this is functional, it is tedious, especially for longer documents.

P1 and P2 mentioned that it was important to be able to retrieve information on-demand about who they were working with and where they were, instead of having it involuntarily pushed to them. In most cases, we observed that participants would skip through them without listening to the entire announcement because they were attending to a different task. On the other hand, blind users have also expressed the need to automatically retrieve real-time updates about where collaborators are located as they are editing to avoid conflicts. For instance, P4 stated that it became 'common courtesy' not to edit a paragraph or sentence if another collaborator's cursor is on it.

6.1.2 Identifying Text Changes. 10 out of 11 participants believed that identifying and tracking text changes in a document is important. P7 said: "I don't want to do anything until I know exactly what it is that I am looking at. I mean, I don't want to write over somebody else's work." None of the participants had a usable strategy for accomplishing this in Google Docs. When possible depending on the context, some participants would read the entire document upon opening it and identifying changes based on their memory of it. This is tedious, as P6 described "It is totally a waste of time reading all those pages just to know where one person had added things, such as a picture or a paragraph."

6.1.3 Working with Comments. 9 out of 11 participants reported that working with comments is important to them when writing collaboratively. P3 mentioned that the commenting feature is "what all of my collaborators make use of to write down their thoughts." P5 also mentioned that commenting is necessary when working on "something hardcore", when "you really need to know why they made the change." P8 and P9 said that they used comments less frequently, and that they prefer to "be clear in the text itself" (P8).

Despite being a commonly desired feature, many participants reported issues with accessing comments in Google Docs. P9 said: "Working with comments is a little bit more lengthy than some other things. It just takes more keyboard shortcuts to access it... Honestly I can't remember them." P1 similarly described the keyboard shortcuts as being too complex, and instead used the easier to remember

shortcut to add a new comment. Then, once their screen reader had focus on the new comment, it was easier to navigate to the comment that they wanted to read. They would later go back and delete the old comment.

In the next few sections, we will report results about participants using CollabAlly to complete collaborative editing tasks.

6.2 Task 1: Locating Collaborators

6.2.1 Task Performance. When the task began, all participants captured CollabAlly's non-speech audio notification of collaborator joining, and most were able to identify and proactively mention that they knew someone had just joined the document. All of the participants were able to identify the correct number of collaborators in the document and their names. They also all identified collaborators' rough locations by reading CollabAlly's prompts such as 'top of page 1.' Participants also used CollabAlly's spatial audio and voice fonts to obtain collaborator locations relative to their current position. Most participants (7 out of 9) were able to identify the direction of the audio and describe the collaborator's locations accordingly. For example: "Ah, she is not as far below me, but still below me" (P3), or "It's coming from the very right, maybe he is sitting at the bottom of the document" (P5). Additionally, most participants (6 out of 9) also successfully used CollabAlly's 'copy selected text to clipboard' feature, pasted the selected text in the search bar, and located exactly where the collaborators were editing in the document.

When P1 and P2 completed the same task with Google Docs, they were not able to obtain such information. Both stated that they were unsure how to locate collaborators, tried navigating through Google Docs' menu bar functions line by line, and tried searching for the keyword "collaborator" in the "help" function in Google Docs, but were unsuccessful. Only one participant in our study (P8) knew about using the collaborators profile pictures at the top right corner of the document, which would read out collaborator names. This information is still limited to names, and requires extra effort of navigating in and out of the document editing area. Note that compared to Google Docs' collaborator announcements, CollabAlly used non-speech audio to prompt users when collaborators joined or left the document, and participants were able to capture all audio prompts when they happen.

6.2.2 Task Feedback. Participants all agreed or strongly agreed that it was easy for them to find collaborators' names and locations with CollabAlly (see Figure 4). P6 said that "it was just a few clicks away and it was easy to access." P8 also mentioned that CollabAlly was

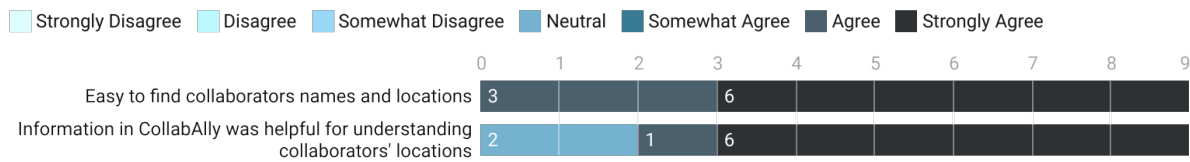


Figure 4: Participants' responses to CollabAlly's collaborator information feature.

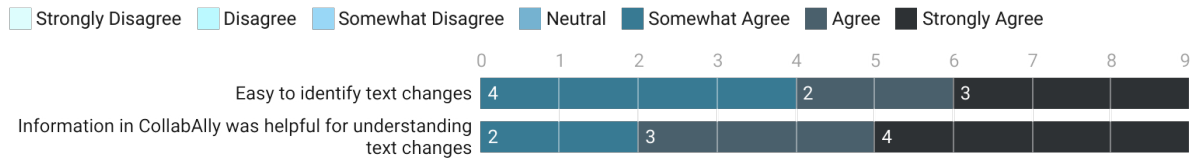


Figure 5: Participants' responses to CollabAlly's text change information feature.

convenient in presenting information in one location, “*compared to before where I needed to scroll up and down in the document.*” Similarly, 7 out of 9 participants agreed or strongly agreed that the collaborator information provided by CollabAlly was helpful for understanding collaborators’ locations. P3 mentioned that the spatial audio “*might help me know whether I need to scroll up or down in the document.*” The other two participants gave a neutral score, as the spatial audio feature gave them inconsistent clues, which caused them some confusion.

Most participants generally agreed that this feature would fit into their current document editing workflows as they do not have an existing strategy for obtaining collaborators’ information. In particular, P5 mentioned that they wanted to be as competitive as sighted collaborators and keep up with multiple people, and CollabAlly gave them a way to do that. Some participants also provided suggestions for improving the collaborator location feature in CollabAlly. For example, P4 wanted to have collaborators’ names integrated and presented to them at the top of the dialog box for them to access as a summary before they navigate through the detailed list.

6.3 Task 2: Identifying Text Changes

6.3.1 Task Performance. All 9 participants were able to successfully identify the document changes, including the change type (whether text or style was added or modified or deleted), author, and its general location. Most participants were able to understand how the text or style was changed by comparing the two presented versions. For example, P8 noted: “*It was ‘comment sensemaking’ before the change, and it’s now ‘comment tracking in the system’*”, and P4 was able to identify that “*the text color was changed, font weight and font size were bigger.*” Participants did not typically jump to the location of the change in the document.

6.3.2 Task Feedback. All participants agreed that it was easy for them to identify text changes with CollabAlly, and that the information provided was helpful in doing so (see Figure 5). However, some (P4, P5, P10) also mentioned that it was harder to identify longer text changes, because they needed to read through them line by line. For example: “*It’s harder to identify longer changes*” (P4),

and “*I can tell what text has been changed, but it’s a little bit tricky to parse because it’s pretty long*” (P5). While Google Docs provides a live-edit tracker and document history page to review changes, as reported in Section 3.2.2, all our participants did not make use of these features because of the difficulty accessing them. With CollabAlly, participants thought that they could be better aware of where text changes happens using the spatial audio features (P7: “*It was hard to locate changes without it (CollabAlly)*”). Furthermore, participants believed that integrating text changes and all other collaborative information in one place not only made collaboration awareness accessible, but also made complex operations simpler. Participants also provided suggestions on improving CollabAlly’s text change features, for example, summarizing the changes in a more readable way because some detected changes included a whole paragraph and required users to manually compare them.

6.4 Task 3: Working with Comments

6.4.1 Task Performance. When the task began, all participants heard CollabAlly’s corresponding non-speech audio notification that a new comment was added. Most participants were able to identify that the sound represented a new comment and began to locate it, though some had to be prompted. Most participants (6 out of 9) were able to navigate to the exact comment locations by copying the selected text and searching it in the document. For the first comment, all participants (either verbally, through a reply, or through a new comment) were able to respond to the question it posed, indicating that they understood it and the context. For the second comment, which asked participants to change two instances of a term in the highlighted paragraph, most participants were able to make the requested changes. Some only located the first instance, indicating that the length of the commented text was unclear.

6.4.2 Task Feedback. All participants agreed or strongly agreed that it was easy to locate and understand comments with CollabAlly (see Figure 6). P6 believed that CollabAlly is “*honestly way better than the methodology that I use*” as their current strategy is to select all the highlighted text and use keyboard shortcuts keystrokes to make or reply to a comment, which is hard to remember and manage. All participants also believed that the information provided by

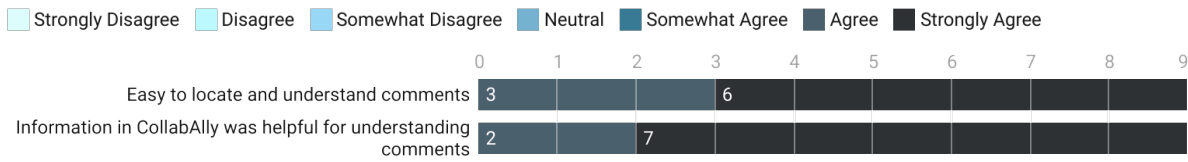


Figure 6: Participants' responses to CollabAlly's comment information feature.

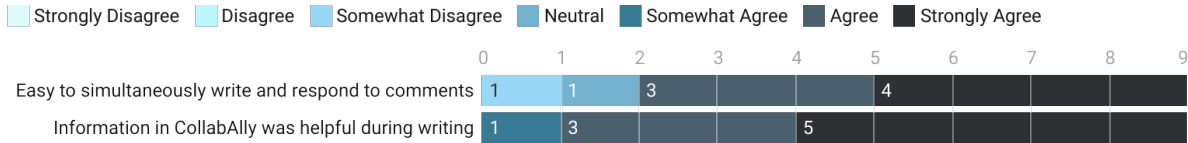


Figure 7: Participants' responses to open usage of CollabAlly for collaborative writing.

CollabAlly was helpful for understanding comments in the document. P4 mentioned the difference between accessing comments with CollabAlly and Microsoft Word. They noted that their screen reader just read out the comment in Microsoft Word, but now “I can choose to open the comment and the text freely because otherwise it will be very cognitively overwhelming.” Some participants (P3, P4, P11) also provided suggestions on making comments ordered or sorted. They wanted to have comments sorted chronologically or in other ways they preferred, like sorting via author.

Compared to Google Docs' commenting features, CollabAlly presents the corresponding document text for each comment (both the highlighted text and the surrounding text on the line) so that blind users can access not only the comment, but also the contextual information. Participants found that this feature enables them to locate context in an efficient way. For example, P5 mentioned that “because it not only presents the position of the comment but also the selected text in the document, so even if it was in the middle of a long document I could find the text or do a quick search for the text.”

6.5 Task 4: Collaborative Writing

6.5.1 Task Performance. Most participants were able to completely respond to the two feedback prompts and address all comments made by the study administrators. Two participants ran out of time to complete the task (but were able to address all of the comments that they read), while two participants left one comment unaddressed each. For the comments that were left unaddressed, they had identical text to an existing comment in a different location, but participants confused them for one another and mistakenly thought that they had already addressed it. Participants used a variety of strategies to write and address comments. Some participants opened CollabAlly to read the comment immediately after hearing the audio notification. Others chose to complete one or both of the prompts completely before addressing comments.

6.5.2 Task Feedback. 7 out of 9 participants felt that it was easy for them to simultaneously write and respond to comments with CollabAlly (see Figure 7). P9 mentioned that having everything in a “centralized location (the dialog box)” made things a lot easier than it would have been without CollabAlly. Most of them liked the background notification sounds and the collaboration awareness

that came with it. P5 believed that it only took a couple of steps to access all the information and the audio was not distracting. Multiple participants mentioned wanting to customize what alerts were on or off for their specific needs, for example P6 mentioned, “I would only keep it for the comments, and turn it off for everything else.” They found the notification for collaborators cursor proximity distracting, saying, “Some people have bad connection, so they would be frequently leaving and that would be really annoying.”

6.6 Overall Feedback

6.6.1 CollabAlly Benefits. From participants' text responses, we summarize the most-liked features of CollabAlly here:

Deep integration with Google Docs. Although collaborative information can be eventually accessed with additional operations in Google Docs, participants preferred not to go through these hurdles. Instead, they felt that CollabAlly consolidated this information, making it easier to access.

Easy navigation. Most participants also mentioned the user interface and usability of CollabAlly. They enjoyed the overall experience with the dialog box and the audio notifications. For example, they particularly liked how the headings were structured in dialog box for better navigation.

Sonification with spatial features. Participants liked how the audio changed when collaboration activities changed in context (P7: “The change in the position of the voices was a big help figuring out where the change was in relation to me”).

Fits into current workflows. Participants also mentioned how CollabAlly would well fit into their current workflow in a non-interrupting way because it “does not interfere during the work, rather it gives me information on request” (P6).

6.6.2 CollabAlly Suggestions. We also summarize the most common suggestions for future improvement:

Audio customization. Many participants suggested methods for customizing CollabAlly's audio, for example, controlling the volume, turn on and off spatial features, change the prompt audio sounds, and customize voice fonts.

Indirect copy and search navigation. Due to a technical limitation of building systems upon Google Docs, CollabAlly provided

indirect navigation by using ‘copy to clipboard’ and asked participants to use the search function after copying. Participants wanted to improve this by having direct navigation: clicking on a button associated with collaborative activity in the document, and directly jumping to its location in context.

Better integration with screen readers. Some users encountered issues with CollabAlly when using JAWS, which moved the screen reader focus back to the document instead of the dialog box.

7 DISCUSSION AND FUTURE WORK

In this section we discuss implications from our research, including discussion on the idea of accessibility versus usability, the possibility of distributing the labor of access, and how we can better integrate with the current assistive technology ecosystem, as was brought up as a major challenge by Das et al. [12]. We also briefly talk about our future deployment and our system’s generalizability.

7.1 Accessibility vs. Usability

Accessibility does not always mean usability. Throughout our iterative co-design process and formal user studies, Google Docs’ version of access is usually not usable. Technically, it provided accessibility features, but it involves a lot of confusing context switches, memorization, cognitive overload, etc. Google Docs is currently a visual-centered system. The information related to collaboration is scattered in different places of the interface, for example, different collaboration activities like comments and text changes have individual windows. Because it is assumed that users can see comments floating next to the document, there is a lack of navigation, such as a ‘jump to this comment’ button. Sighted people can leverage many visual cues to navigate to various locations within the document, like clicking on a collaborator’s avatar to jump to their cursor. For screen reader users, there is no indication that their cursor would move to a new location, making these techniques infeasible.

In our co-design process and studies, one of the most common responses we received when blind users make use of Google Docs features is *“maybe I am not experienced/proficient enough to do this.”* However, this should not have to be a concern for blind users. Similar concerns and findings have been discussed in prior work as well. Li et al. [31] studied the accessibility of various prototyping tools and found that they are largely not accessible via screen readers, including creativity tools like Sketch and Adobe XD. They usually have accessibility issues in input and control instruction, focus order, and keyboard operations. Accessible solutions should follow the original intent of the system, and not just make it technically accessible for users that can spend a significant amount of time to learn the mechanisms only to be missing the benefits that those features should provide.

7.2 Distributing the Labor of Access

Another design implication we drew from our work is to consider the labor of accessibility. What if such a browser extension could also be installed by sighted collaborators? By prompting them to summarize their changes or other collaboration activities, would

this make it easier for blind users to consume? And how will distributing the labor of access potentially improve blind users’ collaboration experience? By proactively summarizing sighted collaborators’ own activities, in the future CollabAlly could not only provide information about who did what at when and where, but the other important factor of why they performed an action. This is motivated by the fact that many participants described the benefits of working synchronously with collaborators. Because sighted collaborators verbally explain their actions and motivations, blind users can make use of that information and thus collaboratively but also independently edit themselves. Future work could investigate how to design systems to encourage such behavior in distributed, asynchronous scenarios, building off of prior work that studied writing behavior in ability diverse teams [12].

7.3 Better Integration with Assistive Technology Ecosystem

We also consider the challenges of developing CollabAlly to work with the ecosystem consisting of tools like Google Docs, Chrome, NVDA, JAWS, VoiceOver, existing keyboard shortcuts, etc. The technical limitation we face is that CollabAlly is built upon these existing systems, but we do not have direct access to information provided by them, only indirect information like public HTML elements. In future work, design of such an add-on system would be easier if developers and designers of different tools collaborate together. For example, CollabAlly’s limitation of indirect navigation could be easily solved by integrating with Google Docs’ page scrolling logic, and CollabAlly’s refocusing issue could also be tackled by working with JAWS to make sure the switch does not drift from dialog box to the document. To this end, we have open sourced our project by documenting how we designed the different layers of CollabAlly, including how we fetched elements from Google Docs, how we parsed them, and how we maintained different states of collaboration activities ([our code repository](#)).

Another limitation of our tool comes from our study design and the length of time the users were given to learn how to use our tool. Although we designed a tutorial session in our formal evaluation study, there will be a learning curve with using any new collaboration system or assistive technology. Ultimately much thought must go in to designing learning materials so that people will adopt the system and a long-term study is needed to show its ultimate usefulness. We recognize this as a potential future work for longitudinal deployments and further iterations of CollabAlly.

7.4 CollabAlly Generalizability

We envision CollabAlly as a base which can be extended in the future towards multiple goals, including different collaborative tasks, different interface representations, and future scalability and evaluations:

Extend Environment Layer. In the future, CollabAlly could be extended beyond document editing to support other forms of digital collaboration. For example, artboards [44], sketching (Sketchboard [48]), designing (InVision [27]), prototyping [31], music composition [18], and programming platforms (Google Colab [22]) could all be enhanced with accessible collaboration awareness.

Extend Representation Layer. As discussed in the system section, CollabAlly has a representation layer that parses visual elements from Google Docs, which can be extended to other representations. For example, if CollabAlly extends to build upon other screen readers, it can incorporate their internal representations to better present its own audio features. CollabAlly can also be extended to model other types of collaboration activities.

Extend Presentation Layer. Similarly, CollabAlly's presentation layer could also be extended with other application-level features, tools, and systems. For example, Das et al. [13] proposed multiple presentations of non-speech audio and text-to-speech voices that CollabAlly integrated. In the future, CollabAlly has the potential in absorbing other research or design's presentations, or integrating other output modalities such as haptics by extending the presentation layer.

Scalability and Evaluations. Although our studies included only three collaborators and a six-page document, CollabAlly's system architecture is designed to be scalable. By providing on-demand access and non-speech audio prompts, CollabAlly could accommodate many more collaborators and collaboration activities. Also, by providing summary of these activities rather than simply providing a list, CollabAlly could work for longer documents and higher-density editing. We hope to evaluate CollabAlly in the future with larger, diverse teams, and in a wider range of editing contexts.

8 CONCLUSIONS

We have presented CollabAlly, a system designed to make collaboration awareness in document editing accessible. Throughout an iterative co-design process, we identified and later reconfirmed challenges of how blind users use Google Docs and complete collaborative writing tasks. We employed a three-layer design to build a prototype system that can make collaboration activities like collaborators' whereabouts, comments, and text changes accessible in an integrated dialog box. CollabAlly also uses audio features like spatial audio, non-speech audio, voice fonts, and text-to-speech to improve the user experience and make contextual information more accessible to blind people. Through a user study with 11 blind participants, we demonstrated that CollabAlly provides improved access to collaboration awareness by centralizing scattered information, sonifying visual information, and simplifying complex operations. CollabAlly also fits well into blind users' current workflows. We then further discussed about the idea of accessibility versus usability, the possibility of distributing the labor of access, how CollabAlly can be better integrated with the current ecosystem, and how CollabAlly can be extended to support broader accessible collaboration in the future.

ACKNOWLEDGMENTS

We thank the participants who contributed to our studies for their time, and the reviewers for their valuable feedback and suggestions. We also thank Robin Brewer, Audrey Labrie, and Sile O'Modhrain for their input and feedback.

REFERENCES

- [1] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: A Navigational Cognitive Assistant

- for the Blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) (*MobileHCI '16*). Association for Computing Machinery, New York, NY, USA, 90–99. <https://doi.org/10.1145/2935334.2935361>
- [2] Khaled Albusays, Stephanie Ludi, and Matt Huenerfauth. 2017. Interviews and Observation of Blind Software Developers at Work to Understand Code Navigation Challenges. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (Baltimore, Maryland, USA) (*ASSETS '17*). Association for Computing Machinery, New York, NY, USA, 91–100. <https://doi.org/10.1145/3132525.3132550>
- [3] Ronald M. Baecker, Dimitrios Nastos, Ilona R. Posner, and Kelly L. Mawby. 1993. The User-Centered Iterative Design of Collaborative Writing Software. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) (*CHI '93*). Association for Computing Machinery, New York, NY, USA, 399–405. <https://doi.org/10.1145/169059.169312>
- [4] Maryam Bandukda and Catherine Holloway. 2020. Audio AR to Support Nature Connectedness in People with Visual Disabilities. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers* (Virtual Event, Mexico) (*UbiComp-ISWC '20*). Association for Computing Machinery, New York, NY, USA, 204–207. <https://doi.org/10.1145/3410530.3414332>
- [5] William A. Bettermann and Timothy Palumbo. 2016. Collaboration Made Easier - Working with Restricted Documents within Office 2013, OneDrive, and Office 365. In *Proceedings of the 2016 ACM SIGUCCS Annual Conference* (Denver, Colorado, USA) (*SIGUCCS '16*). Association for Computing Machinery, New York, NY, USA, 43–45. <https://doi.org/10.1145/2974927.2974959>
- [6] Jeremy Birnholtz and Steven Ibara. 2012. Tracking Changes in Collaborative Writing: Edits, Visibility and Group Maintenance. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) (*CSCW '12*). Association for Computing Machinery, New York, NY, USA, 809–818. <https://doi.org/10.1145/2145204.2145325>
- [7] Meera M Blattner, Denise A Sumikawa, and Robert M Greenberg. 1989. Earcons and icons: Their structure and common design principles. *Human-Computer Interaction* 4, 1 (1989), 11–44.
- [8] BlindSquare. 2021. BlindSquare. <https://www.blindsquare.com>
- [9] Tom Boellstorff, Bonnie Nardi, Celia Pearce, and T. L. Taylor. 2013. Words with Friends: Writing Collaboratively Online. *Interactions* 20, 5 (sep 2013), 58–61. <https://doi.org/10.1145/2501987>
- [10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
- [11] Maria Claudia Buzzi, Marina Buzzi, Barbara Loporini, and Giulio Mori. 2012. Designing e-learning collaborative tools for blind people. *E-Learning-Long-Distance and Lifelong Perspectives* (2012), 125–144.
- [12] 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. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW, Article 191 (nov 2019), 26 pages. <https://doi.org/10.1145/3359293>
- [13] Maitraye Das, Anne Marie Piper, and Darren Gergle. 2022. Design and Evaluation of Accessible Collaborative Writing Techniques for People with Vision Impairments. *ACM transactions on computer-human interaction (TOCHI)* 29, 2, Article 9 (jan 2022), 42 pages. <https://doi.org/10.1145/3480169>
- [14] Paul Dourish and Victoria Bellotti. 1992. Awareness and Coordination in Shared Workspaces. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work* (Toronto, Ontario, Canada) (*CSCW '92*). Association for Computing Machinery, New York, NY, USA, 107–114. <https://doi.org/10.1145/143457.143468>
- [15] Eric Eggert and Shadi Abou-Zahra. 2019. Page structure concepts. <https://www.w3.org/WAI/tutorials/page-structure>
- [16] Thomas Erickson and Wendy A. Kellogg. 2000. Social Translucence: An Approach to Designing Systems That Support Social Processes. *ACM transactions on computer-human interaction (TOCHI)* 7, 1 (mar 2000), 59–83. <https://doi.org/10.1145/344949.345004>
- [17] Robert S. Fish, Robert E. Kraut, and Mary D. P. Leland. 1988. Quilt: A Collaborative Tool for Cooperative Writing. In *Proceedings of the ACM SIGOIS and IEEECS TC-OA 1988 Conference on Office Information Systems* (Palo Alto, California, USA) (*COCS '88*). Association for Computing Machinery, New York, NY, USA, 30–37. <https://doi.org/10.1145/45410.45414>
- [18] Flat. 2022. Online collaborative music notation software. <https://flat.io/en-GB>
- [19] David Goncalves, André Rodrigues, and Tiago Guerreiro. 2020. Playing With Others: Depicting Multiplayer Gaming Experiences of People With Visual Impairments. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) (*ASSETS '20*). Association for Computing Machinery, New York, NY, USA, Article 22, 12 pages. <https://doi.org/10.1145/3373625.3418304>
- [20] Google. 2021. Collaborate & comment with a screen reader. https://support.google.com/docs/answer/6239410?hl=en&ref_topic=6239407
- [21] Google. 2021. Diff-match-patch: repackaging of google's diff match and patch libraries. <https://github.com/diff-match-patch-python/diff-match-patch>

- [22] Google. 2021. Google Colaboratory. https://colab.research.google.com/?utm_source=scs-index
- [23] Google. 2021. Google Docs. <https://www.google.com/docs/about/>
- [24] Google. 2021. Resonance audio. <https://resonance-audio.github.io/resonance-audio/>
- [25] Google. 2021. Text-to-speech. <https://cloud.google.com/text-to-speech>
- [26] Tom Gross, Chris Stary, and Alex Totter. 2005. User-centered awareness in computer-supported cooperative work-systems: Structured embedding of findings from social sciences. *International Journal of Human-Computer Interaction* 18, 3 (2005), 323–360.
- [27] InVision. 2021. InVision | Digital product design, workflow & collaboration. <https://www.invisionapp.com>
- [28] Ida Larsen-Ledet and Marcel Borowski. 2020. “It Looks Like You Don’t Agree”: Idiosyncratic Practices and Preferences in Collaborative Writing. In *32nd Australian Conference on Human-Computer Interaction* (Sydney, NSW, Australia) (OzCHI ’20). Association for Computing Machinery, New York, NY, USA, 339–354. <https://doi.org/10.1145/3441000.3441032>
- [29] Ida Larsen-Ledet and Henrik Korsgaard. 2019. Territorial functioning in collaborative writing. *Computer Supported Cooperative Work (CSCW)* 28, 3 (2019), 391–433.
- [30] 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>
- [31] Junchen Li, Garreth W. Tigwell, and Kristen Shinohara. 2021. Accessibility of High-Fidelity Prototyping Tools. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI ’21). Association for Computing Machinery, New York, NY, USA, Article 493, 17 pages. <https://doi.org/10.1145/3411764.3445520>
- [32] Muhanad S. Manshad, Enrico Pontelli, and Shakir J. Manshad. 2013. Exploring Tangible Collaborative Distance Learning Environments for the Blind and Visually Impaired. In *CHI ’13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) (CHI EA ’13). Association for Computing Machinery, New York, NY, USA, 55–60. <https://doi.org/10.1145/2468356.2468367>
- [33] Daniel Mendes, Sofia Reis, João Guerreiro, and Hugo Nicolau. 2020. Collaborative Tabletops for Blind People: The Effect of Auditory Design on Workspace Awareness. *Proceedings of the ACM on Human-Computer Interaction* 4, ISS, Article 197 (nov 2020), 19 pages. <https://doi.org/10.1145/3427325>
- [34] Microsoft. 2021. Microsoft Word. <https://www.microsoft.com/en-us/microsoft-365/word>
- [35] Microsoft. 2021. Use a screen reader to add, read, and delete comments in Word. <https://support.microsoft.com/en-us/office/use-a-screen-reader-to-add-read-and-delete-comments-in-word-c4648d50-4f7b-4795-b487-9e3141a5a596#PickTab=Windows>
- [36] Microsoft. 2021. Use a screen reader to track and review changes in a document in Word. <https://support.microsoft.com/en-us/office/use-a-screen-reader-to-track-and-review-changes-in-a-document-in-word-8d415281-6ef2-41ea-8532-38e410be5988#PickTab=Windows>
- [37] Microsoft Research. 2021. Microsoft Soundscape – A map delivered in 3D sound. <https://www.microsoft.com/en-us/research/product/soundscape/>
- [38] Jonas Moll and Eva-Lotta Sallnäs Pysander. 2013. A Haptic Tool for Group Work on Geometrical Concepts Engaging Blind and Sighted Pupils. *ACM Transactions on Accessible Computing (TACCESS)* 4, 4, Article 14 (jul 2013), 37 pages. <https://doi.org/10.1145/2493171.2493172>
- [39] Sylvie Noël and Jean-Marc Robert. 2004. Empirical study on collaborative writing: What do co-authors do, use, and like? *Computer Supported Cooperative Work (CSCW)* 13, 1 (2004), 63–89.
- [40] Nuance. 2021. Nuance Vocalizer. <https://www.nuance.com/omni-channel-customer-engagement/voice-and-ivr/text-to-speech/vocalizer.html>
- [41] Ricardo Olenewa, Gary M. Olson, Judith S. Olson, and Daniel M. Russell. 2017. Now That We Can Write Simultaneously, How Do We Use That to Our Advantage? *Commun. ACM* 60, 8 (jul 2017), 36–43. <https://doi.org/10.1145/2983527>
- [42] Judith S. Olson, Gary M. Olson, Marianne Storösten, and Mark Carter. 1993. Groupwork Close up: A Comparison of the Group Design Process with and without a Simple Group Editor. *ACM Transactions on Information Systems (TOIS)* 11, 4 (oct 1993), 321–348. <https://doi.org/10.1145/159764.159763>
- [43] Judith S. Olson, Dakuo Wang, Gary M. Olson, and Jingwen Zhang. 2017. How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 1, Article 4 (mar 2017), 40 pages. <https://doi.org/10.1145/3038919>
- [44] Anastasia Schaadhardt, Alexis Hiniker, and Jacob O. Wobbrock. 2021. Understanding Blind Screen-Reader Users’ Experiences of Digital Artboards. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI ’21). Association for Computing Machinery, New York, NY, USA, Article 270, 19 pages. <https://doi.org/10.1145/3411764.3445242>
- [45] John G. Schoeberlein and Yuanqiong Wang. 2012. Accessible Collaborative Writing for Persons Who Are Blind: A Usability Study. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility* (Boulder, Colorado, USA) (ASSETS ’12). Association for Computing Machinery, New York, NY, USA, 267–268. <https://doi.org/10.1145/2384916.2384986>
- [46] John G. Schoeberlein and Yuanqiong Wang. 2013. Improving Collaborative Writing for the Blind. *ACM SIGACCESS Accessibility and Computing* 107 (sep 2013), 11–16. <https://doi.org/10.1145/2535803.2535805>
- [47] John Gerard Schoeberlein and Yuanqiong Wang. 2014. Usability Evaluation of an Accessible Collaborative Writing Prototype for Blind Users. *Journal of Usability Studies* 10, 1 (nov 2014), 26–45.
- [48] Sketchboard. 2021. Online collaborative whiteboard | sketchboard. <https://sketchboard.io/>
- [49] John Tang. 2021. Understanding the Telework Experience of People with Disabilities. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1, Article 30 (apr 2021), 27 pages. <https://doi.org/10.1145/3449104>
- [50] Dakuo Wang, Judith S. Olson, Jingwen Zhang, Trung Nguyen, and Gary M. Olson. 2015. DocuViz: Visualizing Collaborative Writing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI ’15). Association for Computing Machinery, New York, NY, USA, 1865–1874. <https://doi.org/10.1145/2702123.2702517>
- [51] Dakuo Wang, Judith S Olson, Jingwen Zhang, Trung Nguyen, and Gary M Olson. 2015. How students collaboratively write using google docs. *ICConference 2015 Proceedings* (2015).
- [52] Dakuo Wang, Haodan Tan, and Tun Lu. 2017. Why Users Do Not Want to Write Together When They Are Writing Together: Users’ Rationales for Today’s Collaborative Writing Practices. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW, Article 107 (Dec. 2017), 18 pages. <https://doi.org/10.1145/1313742>
- [53] Mirza Muhammad Waqar, Nafees Ahmad, Muhammad Aslam, and AM Martinez-Enriquez. 2020. Cloud based Co-authoring platform for visually impaired people. In *2020 17th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. IEEE, 1–6.
- [54] 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). <https://doi.org/10.3390/sym11020238>
- [55] Ryan Wedoff, Lindsay Ball, Amelia Wang, Yi Xuan Khoo, Lauren Lieberman, and Kyle Rector. 2019. Virtual Showdown: An Accessible Virtual Reality Game with Scaffolds for Youth with Visual Impairments. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (CHI ’19). Association for Computing Machinery, New York, NY, USA, 1–15. <https://doi.org/10.1145/3290605.3300371>
- [56] Gareth R. White, Geraldine Fitzpatrick, and Graham McAllister. 2008. Toward Accessible 3D Virtual Environments for the Blind and Visually Impaired. In *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts* (Athens, Greece) (DIMEA ’08). Association for Computing Machinery, New York, NY, USA, 134–141. <https://doi.org/10.1145/1413634.1413663>
- [57] Soobin Yim, Dakuo Wang, Judith Olson, Viet Vu, and Mark Warschauer. 2017. Synchronous Collaborative Writing in the Classroom: Undergraduates’ Collaboration Practices and Their Impact on Writing Style, Quality, and Quantity. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW ’17). Association for Computing Machinery, New York, NY, USA, 468–479. <https://doi.org/10.1145/2998181.2998356>