

The Sky is the Limit: Understanding How Generative AI can Enhance Screen Reader Users' Experience with Productivity Applications

Minoli Perera

Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
minoli.perera@monash.edu

Cagatay Goncu

Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
cagatay.goncu@monash.edu

Swamy Ananthanarayan

Department of Human-Centred Computing
Monash University
Melbourne, Victoria, Australia
swamy.ananthanarayan@monash.edu

Kim Marriott

Department of Human-Centred Computing
Monash University
Melbourne, Australia
kim.marriott@monash.edu

Abstract

Productivity applications including word processors, spreadsheets, and presentation tools are crucial in work, education, and personal settings. Blind users typically access these tools via screen readers (SRs) and face significant accessibility and usability challenges. Recent advancements in Generative AI (GenAI) may address these challenges by enabling natural language interactions and contextual task understanding. However, there is limited understanding of SR users' needs and attitudes toward GenAI assistance in these applications. We surveyed 99 SR users to gain a holistic understanding of the challenges they face when using productivity applications, the impact of these challenges on their productivity and independence, and their initial perceptions of AI assistance. Driven by their enthusiasm, we conducted interviews with 16 SR users to explore their attitudes toward GenAI and its potential usefulness in productivity applications. Our findings highlight its need to support existing SR workflows and the importance of enabling customization and task verification.

CCS Concepts

• **Human-centered computing** → **Empirical studies in accessibility.**

Keywords

blind, accessibility, productivity applications, assistive technology, screen readers, AI assistants, Generative AI, virtual assistants

ACM Reference Format:

Minoli Perera, Swamy Ananthanarayan, Cagatay Goncu, and Kim Marriott. 2025. The Sky is the Limit: Understanding How Generative AI can Enhance Screen Reader Users' Experience with Productivity Applications. In *CHI '25, Yokohama, Japan*



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

CHI '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1394-1/25/04

<https://doi.org/10.1145/3706598.3713634>

Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3706598.3713634>

1 Introduction

Productivity applications play an integral part in current workplace, higher education and in personal settings. These tools are mainly included in productivity suites such as Microsoft 365 [78], Google Workspace [50] and Apple iWork [13]. Although they are widely used around the world and have become a necessity to perform numerous tasks [67], the user interfaces for most of these applications are predominantly designed for sighted user interactions and heavily depend on visual indicators [61, 65]. As a result, blind users often require more time and effort to learn and complete tasks compared to their sighted coworkers [21, 89, 106].

Blind productivity application users primarily rely on screen readers (SRs) to interact with these tools. Unfortunately, the inherent non-visual interaction and reliance on auditory feedback of SRs present numerous accessibility and usability challenges for their users, such as when understanding visual information [56, 99], obtaining overview information [109, 110], finding content [84] or working collaboratively [17, 36]. While these challenges have been examined in individual applications, they have not been holistically explored across the full spectrum of productivity applications to understand the challenges most affecting blind users and to prioritize areas that need to be addressed.

Prior research has shown that virtual assistants such as Siri [14], Alexa [12] and Google Assistant [49] are beneficial for blind people [1, 2, 22, 94, 98], but are limited to lightweight tasks such as playing music, determining weather [71, 73], accessing emails, and setting reminders [2]. Recently introduced Generative AI (GenAI) tools based on large language models (LLMs) [24, 60] have further extended the capabilities of virtual assistants by enabling more advanced tasks with contextual understanding and sophisticated natural language interactions. These tools have gained immense popularity and have been incorporated into various domains, including accessibility. Such tools are available through web interfaces (e.g., OpenAI ChatGPT [88]), SRs (e.g., JAWS Picture Smart

AI [102]), mobile and desktop applications (e.g., Be my AI [20], Microsoft Copilot [79]), and have been found to improve accessibility for blind people [6, 9, 46, 47]. Prior research in this area has primarily focused on image-based visual assistance tasks [9, 33, 47]. While these tools show promise for improving the experiences of SR users [44, 61, 112], their use in productivity-related tasks remains notably underexplored.

In this paper, we explore the broader implications of the accessibility challenges faced by blind productivity applications users and the potential role of GenAI in addressing these challenges. Through an online survey involving 99 SR users and semi-structured interviews with 16 SR users, we offer contributions in the following areas:

- Holistic understanding of the accessibility issues faced by SR users in productivity applications (without GenAI) and their impact on productivity and independence.
- An understanding of the attitudes of SR users towards the usefulness of GenAI when working with productivity applications and their attitudes to trust, privacy and other potential concerns.
- An understanding of how SR users want to integrate and interact with GenAI when working with productivity applications.

The outcomes presented in this paper will provide valuable insights for productivity application and assistive technology designers, enabling them to understand the challenges and accessibility needs of SR users, as well as the potential of GenAI in enhancing SR experience with productivity applications.

2 Related Work

2.1 Accessibility and Usability of Productivity Applications

Productivity applications are widely and regularly used in work and education settings. A survey with 308 people who are blind or have low vision (BLV) employed in the U.S. revealed that around 95% used some type of productivity applications during a typical workweek [106]. Blind users primarily use SRs with synthetic speech to access these applications. The linear, temporal nature of auditory screen reader feedback imposes high cognitive demands, making it challenging to access application information [109]. In spreadsheets, cell-by-cell navigation obscures data relationships and trends [57, 92]. Likewise, the lack of an overview in presentation tools complicates the understanding of slide layouts and object relationships [99]. Previous studies have highlighted that SR users face challenges in accessing graphical content, such as diagrams and images, primarily due to the lack of alternative descriptions [55, 106, 120] and the inability of SRs to convey spatial information [56, 58]. Shared screens on video conferencing platforms remain largely inaccessible to SR users [28, 29, 68, 69, 106]. Understanding and applying document formatting is a common challenge for SR users, often requiring sighted assistance [41, 83]. Users struggle with elements like indentation, spacing, and font styles [41]. SRs either provide overly verbose details or omit essential information (e.g., spacing between words), resulting in formatting errors and

inconsistencies [83]. In document collaboration, SR users face challenges with tracking changes, comments [26], and real-time editing notifications [36]. The linear nature of SR output makes it difficult to distinguish between document content and collaboration markup, understand the context of edits, and track collaborator contributions [34, 36, 100].

Furthermore, SR users encounter issues finding and accessing features due to the need to navigate through menus and toolbars [25, 66, 99, 106] or because of unlabelled controls [106]. SR-related issues such as the lack of keyboard commands to perform actions [25] and SRs not responding [25], as well as inconsistencies among different types of SRs, have also been reported across various applications [34]. Prior work has also found SR users often need to seek sighted assistance to access information [110] and create or manipulate content [99] due to accessibility issues within these applications. A limitation of prior research is that research into these issues have focused on individual applications, and we do not have a holistic understanding of the challenges that SR users face across the suite of productivity applications or the impact that this has on them.

Several studies have aimed to improve the accessibility and usability of productivity applications for blind users by introducing feature enhancements [64, 91], interface customization [37, 64, 84, 90, 100, 114], auditory feedback [64, 109, 110], and multi-modal and interactive interfaces [35, 38, 117, 121, 122]. However, there has been limited investigation into how AI technologies can be leveraged to address the unique challenges faced by SR users in these applications.

2.2 AI Assistants

Virtual assistants such as Siri [14], Alexa [12], and Google Assistant [49] are popular among blind and sighted users, primarily for simple tasks like playing music or checking the weather [71, 73], with limited use in productivity tasks such as accessing emails or scheduling meetings [2]. Despite challenges such as speech recognition errors and limited task follow-up capabilities, prior studies [2, 22, 94] highlight their benefits, offering more efficient and natural interactions compared to SRs or touch gestures [2].

Recently introduced GenAI tools have extended the capabilities of virtual assistants, enabling contextual task understanding and more advanced natural language interactions. GenAI tools, such as ChatGPT [88], Google Gemini (formerly Bard) [48], and Copilot [79], have gained immense popularity and have been incorporated into various domains, including work [27, 40], education [8, 75], programming [39, 116], content creation [54], and accessibility [63, 111]. Considering its application in different domains, Weisz et al. [118] have introduced design principles that aid design practitioners, highlighting the importance of designing for generative variability, co-creation, and imperfections. GenAI tools have recently been integrated into the technology stack to improve productivity and accessibility. These tools are now incorporated into productivity applications (e.g., Copilot [79]), screen readers (e.g., JAWS Picture Smart AI [102]), desktop applications (e.g., Be My Eyes for Windows [19], Savant [61]) and mobile applications (e.g., Seeing AI [80], Be My AI [20], Aira [7]). They allow users to interact with applications or content (image or text) using natural language commands, extracting information and performing tasks.

Recent studies [6, 9, 47], have explored how blind users understand and use GenAI tools. Adnin and Das [6] found that blind users often develop flawed or oversimplified mental models of GenAI, influenced by their experiences with blindness, use of assistive technologies, and the accessibility of GenAI tools. Blind users have also encountered accessibility issues when using these tools with their SRs, particularly due to unlabeled and untagged content [6, 11]. Alharbi et al. [9] reported that blind users employ various verification methods to address GenAI inaccuracies and hallucinations [30, 108], including experimentation, sighted assistance, and cross-referencing with different devices and applications. Additionally, prior research [6, 9, 46] highlights that these tools often exhibit negative biases towards blind people.

Prior research in this area has primarily focused on image based visual assistance [6, 9, 33, 47]. While recent studies [44, 61, 112] have highlighted the potential of GenAI tools to enhance SR users' experiences in productivity applications via conversational interactions, there has been limited investigation into how these technologies can be leveraged to address the unique challenges encountered by SR users in these applications. Our work serves as an initial step in understanding SR users' expectations regarding the capabilities of GenAI tools in addressing these challenges, their preferred methods for using and integrating these tools with SRs, and their concerns about working with them in productivity applications.

3 Survey

We conducted an online survey to gain a holistic understanding of the accessibility issues faced by SR users with productivity applications, and SR users initial perceptions of the usefulness of AI assistance in this context. The survey was created using Google Forms, which was tested for accessibility using SRs such as JAWS [101], NVDA [5], VoiceOver [53] and Microsoft Narrator [81]. As the first step of the survey, participants were presented with consent details and asked to provide electronic consent before proceeding to the questions. The survey consisted of 21 main questions including demographics ($n=6$), multiple-choice with the option to add further responses ($n=7$), Likert-scale ($n=4$), and open-ended questions ($n=4$).

In the first half of the survey, we collected users' demographic information, assessed their comfort level with technology, explored their use of productivity applications, and inquired their learning methods. The applications were derived from an examination of existing productivity suites [13, 50, 78]. Then, we posed open-ended questions to elicit information about the worst accessibility issues and the applications in which these issues occur. We also inquired about the impact of these issues on users, their feelings when encountering these problems and the strategies employed to overcome them. Lastly, we explored users initial perceptions of the usefulness of AI assistance in productivity applications. Since terms like GenAI were not widely recognized among the general public at the time the survey was conducted (January–April 2023), we described AI assistance as a virtual assistant capable of voice- and chat-based interactions.

The survey was promoted online through mailing lists dedicated to individuals who are blind (NFB BlindResearch [86], VIP-1 [42], ATUG [51], BCA newsletter [16] and SR user groups [4]). Ten randomly selected respondents from the respondent pool were given

\$30 Amazon gift cards. The study was approved by our institutional ethics committee.

3.1 Participants

A total of 102 individuals responded to the survey. During the refinement process, it was identified that one response was invalid, and two respondents did not use SRs with productivity applications. The data reported herein is from the remaining 99 valid responses. The majority of respondents (68%) reported being totally blind, while 32% reported being legally blind. In terms of gender, 55% identified as women, 43% as men, and 2% as non-binary. Most participants resided in the United States (77%), followed by 14% from Australia. The remaining participants were from other countries, including Canada (2%), India (2%), New Zealand (2%), the United Kingdom (1%), Germany (1%), and Bangladesh (1%). Among the respondents, 18% were aged between 18 to 34, 54% aged between 35 and 54, and the remaining participants were 55 years old or older. Regarding educational qualifications, 75% of participants held a bachelor's degree or higher. In terms of comfort level with technology, the majority (78%) reported being up-to-date with current technology, 20% identified as early adopters, and 13% expressed a need for support or encouragement to use new technologies.

3.2 Data Analysis

Responses for open-ended questions were thematically analyzed using an inductive approach [23]. The first author conducted open coding of all responses, using low-level descriptive codes. These initial codes were then collaboratively refined with another co-author to develop high-level codes. For example, low-level codes such as "difficulties understanding shared screens" and "difficulties understanding images" were grouped under the high-level code "difficulties understanding graphical content." All authors met weekly to review these low-level codes, their associated excerpts, and the formation of high-level codes, refining the coding process and mitigating individual biases. Through this iterative process, we identified six themes related to accessibility issues, further detailed in Section 3.3.2.

3.3 Findings

3.3.1 Use of Productivity Applications and Methods of Access. Table 1 shows the most regularly used productivity applications (both desktop and web). Productivity applications were primarily used in personal (97%) and work settings (75%), with some using it for educational (32%) and other purposes such as volunteer work (6%). While all used SRs, more than one third of them (39%) said that they had to ask sighted people for assistance when accessing applications. Additionally, some also used braille displays (36%), remote assistance services such as Aira or Be My Eyes (26%) and magnifiers (6%) to access applications.

3.3.2 Accessibility and Usability Challenges Faced by Screen Reader Users. The following themes emerged from responses to the open-ended question related to the worst accessibility issues: understanding content, accessing controls, authoring and editing, collaboration, navigation, and SR related issues. Table 2 illustrates the number of respondents experiencing issues in each of these themes. The

Table 1: Most regularly used productivity applications.

Application	(%)
Email (e.g. Microsoft Outlook, Gmail etc.)	98
Video conferencing (e.g. Microsoft Teams, Zoom, Google Meet etc.)	98
Word processing (e.g. Microsoft Word, Google Docs, Apple Pages etc.)	96
Spreadsheets (e.g. Microsoft Excel, Google Sheets, Numbers etc.)	85
Calendar (e.g. Microsoft Outlook, Google Calendar etc.)	84
Cloud storage (e.g. Microsoft OneDrive, Google Drive etc.)	78
Instant messaging (e.g. Microsoft Teams, Google Chat, Slack etc.)	62
Presentation tools (e.g. Microsoft PowerPoint, Google Slides, Keynote etc.)	49
Note taking tools (e.g. Microsoft OneNote, Google Keep etc.)	28
Other (Database tools, Coding platforms etc.)	7
Digital whiteboard tools (e.g. Google Jamboard, Miro etc.)	4

applications with the most issues were whiteboard tools (user ratio=1) and presentation tools (0.86), followed by word processing applications (0.53) and spreadsheets (0.46). Although applications such as email, video conferencing, calendar, and instant messaging are heavily used, fewer issues were reported in them. The majority of users experienced issues in accessing content and controls. A more detailed breakdown of issues is listed in Appendix A.

(1) Understanding Content (n=57): The majority of users (n=35) reported the inability to understand graphical data in emails, documents, spreadsheets, and presentations, primarily attributed to the lack of alternative text. Users encountered difficulties in understanding charts, scanned documents, screenshots, video content, and icons, as well as accessing shared screens in video conferencing platforms. Users (n=26) also encountered challenges in understanding non-graphical data, such as links, tabular data, and formulas. The use of color and proximity as the sole means to convey information posed challenges in extracting information, *“sometimes, co-workers will mark important cells by changing the background or text color. In large spreadsheets, it is difficult to find the marked cells [...] It can be time consuming to manually query each cell’s appearance to find the ones that are different”*. Users (n=10) also faced difficulties in accessing different styles of bullet lists and numbering, as well as detecting formatting issues such as inconsistencies in font size or style within documents.

(2) Accessing Controls (n=41): Due to unlabelled elements in the applications, users (n=31) were unable to understand their purpose or access the features associated with them. These elements primarily included unlabelled form fields, drop-downs and buttons, *“A button will be not labelled properly, so it just says button”*. Users (n=7) commented on ribbon menus as being *“slow and annoying”* and *“hard to navigate”*. Users (n=8) faced difficulties due to non-functional keystrokes or struggled to remember keystrokes to perform actions due to *“too many keyboard shortcuts that differ from app to app”*.

(3) Authoring and Editing (n=28): When creating content, users (n=10) encountered several challenges, particularly within more visually-oriented tools like presentation and white-boarding applications. Users faced difficulties in grasping the visual arrangement, managing spacing, and adjusting the position or size of elements to align with the content, *“Not being able to insert graphics/photos/other features like bullet points or resize those first two while keeping text*

where it belongs on the page. No confidence in where items are on the final product”. Users (n=6) also struggled with formatting content, *“It is difficult to tell if the column width in Excel cuts off the data from a sighted viewer”*. These challenges led users (n=3) to seek assistance from sighted individuals, *“At the moment, human assistance is really the only option to be 100% sure that a presentation looks nice”*. Users (n=11) encountered challenges in accurately identifying the appropriate location for implementing edits, *“It will appear that I am in one part of the document, but when I attempt to do an edit, the text is inserted into another part of the document”*. Some users (n=6) were frustrated with spell check and auto-correct features, *“Revision, editing, and auto correction tools are poorly designed. Now even running a spelling check, creates anxiety in me”*.

(4) Navigation (n=18): Due to poor content structure users (n=13) experienced difficulties navigating within documents. Untagged content, for instance *“lack of the use of headings where appropriate”*, negatively affected the reading order in documents. One user noted that *“things are labelled poorly so the screen reader jumps around, reading paragraph text before headers”*. Beyond large documents, users also found navigation within spreadsheets and presentations to be particularly challenging. Additionally, users (n=5) reported complexities in *“navigating around various components”* of application interfaces. For instance, one participant highlighted difficulties transitioning between the chat view and the files view in Microsoft Teams.

(5) Screen Reader (n=18): Instances where SRs did not respond hindered users’ (n=12) access to applications, *“I can’t perform a needed function with the screen reader because it doesn’t want to work”*. Users (n=4) reported difficulties with SR losing focus and jumping to random parts of the screen.

(6) Collaboration (n=16): While collaborating with others, users (n=9) encountered challenges in understanding modifications made by their peers. One user explained, *“making sense of track changes in Word. Inability to do this makes it impossible to collaborate on documents without creating extra work for others”*. Additionally, one user mentioned that real-time collaboration on documents led to decreased responsiveness in SRs, resulting in what was deemed a *“frustratingly inefficient experience”*. Users experienced uncertainty when making edits to documents, as they were unsure whether the changes would affect the work of others.

Table 2: No. of users who reported issues in different productivity applications.

Themes	Email	Word processing	Video conferencing	Spreadsheets	Calendar	Cloud storage	Instant messaging	Presentation tools	Note taking tools	Whiteboard tools	Occurs in all apps	Total
High level general comments about inaccessible apps	4	9	4	5	2	1	-	12	-	2	1	26
(1) Understanding Content	9	25	8	19	1	2	-	20	1	1	10	57
(2) Accessing Controls	2	13	12	10	2	3	3	10	-	-	13	41
(3) Authoring and Editing	1	16	-	6	2	-	-	11	2	1	4	28
(4) Navigation	1	11	4	5	-	3	2	5	-	-	1	18
(5) SR-related Issues	3	6	4	4	-	1	-	4	1	-	3	18
(6) Collaboration	-	11	2	5	-	-	-	1	-	1	-	16
Total No. of users who reported issues	12	51	25	39	5	7	3	42	3	4	26	92
User Ratio*	0.12	0.53	0.26	0.46	0.06	0.09	0.05	0.86	0.11	1	-	-

*User Ratio = Total no. of users reporting issues/Total no. of users of applications.

3.3.3 Learning and Troubleshooting Strategies. The most popular ways respondents learned to use productivity applications was through self-teaching (80%), from other individuals who are blind (69%) and from application/SR companies support platforms/training materials (63%). Some respondents (43%) said they learn from sighted individuals. The main difficulty users faced when self-learning was the inaccessibility of training materials ($n=14$), which had instructions based on visuals and often focused on sighted users. One user commented, “there are also lots of videos where they say things like, “Click here” or don’t provide enough context to be useful. Similarly some training materials rely heavily on screenshots or graphics to convey essential information”. Additionally, users ($n=9$) found it hard to “find accessible and screen reader/keyboard oriented instructions” and new features lacked SR-focused documentation. Users’ ($n=15$) reported having a “steep learning curve” and having to put “extra time to learn compared to a sighted person.”

The most common strategies blind users typically use when encountering difficulties with productivity applications were seeking assistance from a sighted person (84%) or a blind person (81%). Respondents also reported using Google search (76%), remote assistance services such as Aira, Be My Eyes (49%), and AI chatbots such as ChatGPT (8%). Almost one third of the users (34%) reported skipping tasks due to difficulties in solving issues.

3.3.4 Impact on Productivity and Independence. Table 3 shows the impact of accessibility issues on users’ productivity and independence ($N^1=92$), along with their emotional responses to these issues ($N=91$). Users reported that they need to invest more time and effort than sighted users to complete tasks, “I often have to spend a lot more time outside work to keep up and meet deadlines that others seem to do during the work day”. Participants indicated that accessibility issues in productivity applications had a significant impact on their employment opportunities, “it means that often work opportunities are reduced or removed. It means that often I cannot use computer systems, getting mentally and emotionally frustrated, so I give up and miss out on normal tasks and social interactions in life and in the workplace”. Another user said, “assistive technology programs

having too many compatibility issues with productivity applications used by my former employer/company is the main reason I’m no longer employed”. We identified that accessibility issues also caused reduced confidence and distress in users. One user shared that, “it is psychologically hard, daunting and depressing to bang your head against those artificially created walls”.

3.3.5 Initial Exploration of the Usefulness of AI Assistance in Productivity Applications. Figure 1 illustrates the users’ perceived usefulness of AI assistance in specific aspects of productivity applications. On average, about 75% of respondents considered AI assistance to be extremely to moderately useful for answering specific questions about data, performing tasks, providing an overview and assisting with learning. Users preferred the technology to assist with tasks such as describing visual data, creating visually appealing content, providing help with understanding and formatting documents, and performing advanced calendar operations. However, four users commented that they did not perceive the need for an AI assistance if all content were accessible. Additionally, four users preferred human assistance over AI assistance.

3.4 Survey Discussion

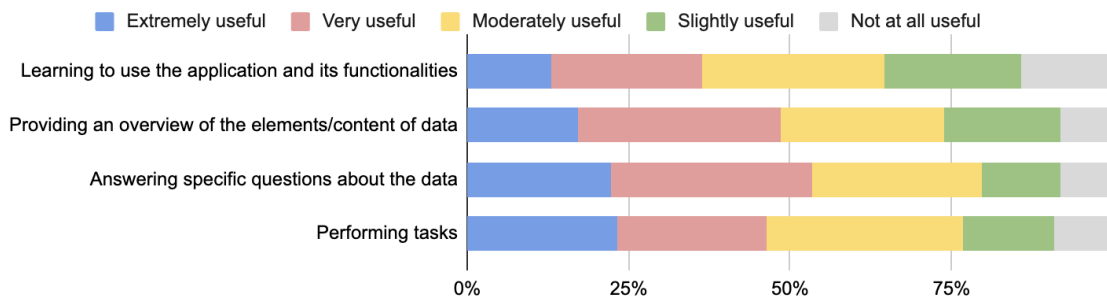
Our survey confirmed that the challenges identified in previous research continue to hold for spreadsheets [57, 106], word processing [34, 82], presentation tools [99] and video conferencing platforms [68, 69]. In comparison with applications such as presentation tools, word processors, and spreadsheets, where more than half of their users reported issues, email, video conferencing, calendar, and instant messaging incurred less issues for SR users. Although presentation tools are widely used in higher education and workplace settings across the general population [99], our survey showed a lack of usage among SR users, with only 49% reporting their use. Furthermore, digital white-boarding tools shows significantly less usage (4%), with all its users reporting accessibility issues. This accords with prior work [99].

Our survey also underscored that accessibility and usability issues in applications have an alarming adverse impact on productivity, information extraction, task performance, and independence

¹ N represents the number of respondents who answered the corresponding question.

Table 3: Ways in which accessibility issues regularly affect SR users.

Impact	(%)
It takes me longer than my sighted colleagues to complete tasks	91
I miss important information which is conveyed through visual cues	84
I am unable to perform some tasks	75
I rely on someone else's assistance to complete some tasks	68
Other (Stress and frustration)	4
Emotional Response	(%)
I feel frustrated because some features are not accessible to me	92
I dislike feeling dependent on help from colleagues	77
I feel less productive than my sighted colleagues	74
I feel excluded from some activities	71
I feel less confidence	52
I feel less useful than my sighted colleagues	52
Other (Grumpy, irritable, depressed, insecure and anxious)	2

**Figure 1: Usefulness of AI assistance in productivity applications.**

(Table 3), and that SR users often require sighted assistance to access, learn, and overcome difficulties (Sections 3.3.1 and 3.3.3). It reveals that these issues result in emotional distress, including frustration, exclusion, reduced confidence, and diminished self-worth for SR users (Table 3). While prior studies [34, 99, 106] have identified negative experiences among SR users in these applications, our study provides further compelling evidence that accessibility issues in productivity applications continues to gravely affect a large number of SR users.

Our study also revealed that SR users believe AI assistance could be beneficial when working with productivity applications (Section 3.3.5). However, the survey did not explore in depth how they thought AI assistance could address their challenges or how they envisioned using it. We explore this in our second study.

4 Interview Study

Motivated by the positive attitudes toward AI assistance in productivity applications (Section 3.3.5), we conducted semi-structured interviews with 16 SR users to further explore how best to apply AI assistance in this context. Given the widespread popularity and rapid advances in GenAI at the time of the interviews (mid-2024), the study focused on exploring its potential use. The study aimed to address the following open questions: (1) What is the perceived usefulness of GenAI assistance when working with productivity applications? (2) How do SR users prefer to interact with GenAI in productivity applications? (3) What potential concerns arise when using GenAI in productivity applications?

4.1 Participants

We recruited 16 participants (11 totally blind, 5 legally blind; 8 male, 8 female) who use SRs to access productivity applications. Fifteen participants were recruited from the survey, having shown interest in participating in further studies, and one participant was recruited through professional contacts. Participant ages ranged from 18 to 64. Table 4 provides a summary of participant demographics. All participants had prior experience using GenAI tools (e.g., Be My AI, Picture Smart AI, Seeing AI). They currently used these tools to describe photos, charts or screenshots ($n=13$) and reading printed documents ($n=4$). Some participants also used tools like ChatGPT and Copilot for writing tasks ($n=5$), summarizing content ($n=4$), generating code ($n=2$) and gaining general knowledge ($n=6$) as well as learning about academic concepts ($n=2$). We deliberately recruited participants who had revealed varying opinions about usefulness of AI assistance in the survey to capture a diverse range of perspectives and better understand the reasons behind their preferences.

4.2 Procedure

We conducted semi-structured interviews via Zoom, all of which were audio-recorded. Interviews lasted between 60 and 90 minutes, with an average duration of 80 minutes (Participants had previously received information about the project and given consent via email). During the interviews, we began by asking participants to discuss their current use of virtual assistants such as Siri, Alexa, Google

Table 4: Details of study participants.

P#	Vision Level	Age	Country	Occupation	Virtual assistants used	GenAI tools used/tried	Screen Reader
P1	Totally blind	55-64	Australia	Assistive Technology Centre Manager	Siri, Alexa, Google Assistant	Aira, Picture Smart AI, Copilot on PC	JAWS
P2	Totally blind	18-24	India	Research Assistant and Assistive Technology Trainer	Cortana, Google Assistant	ChatGPT, Gemini, Copilot on Web, Be my AI on PC	NVDA
P3	Totally blind	35-44	Singapore	IT Executive	Siri, Alexa, Google Assistant	ChatGPT	JAWS, NVDA
P4	Totally blind	45-54	United States	Volunteer at Blind People Focused Organizations	Siri, Cortana	Bard, Be my AI, Seeing AI	JAWS
P5	Totally blind	45-54	United States	Cyber Security Student	Siri, Alexa, Google Assistant, Cortana	ChatGPT, Be my AI	JAWS, NVDA
P6	Legally blind	55-64	United States	Self-publishing Author and Toy Maker	Siri, Cortana	Be my AI, PiccyBot [15], Seeing AI	JAWS, NVDA, Narrator
P7	Totally blind	25-34	United States	Assistive Technology Trainer	Cortana, Siri, Google Assistant, Bixby [97]	NVDA ChatGPT addon, ChatGPT, Be my AI, Ollama [87]	NVDA, JAWS
P8	Legally blind	45-54	Australia	Customer Service Representative	Siri	ChatGPT, Copilot on PC, Meta AI [77], Seeing AI	Narrator
P9	Totally blind	55-64	United States	Account Manager	Siri, Alexa, Google Assistant	Be My AI, Picture Smart AI	JAWS
P10	Totally blind	45-54	Canada	User Experience Accessibility Consultant	Google Assistant, Alexa, Cortana, Siri	Be My AI, Picture Smart AI, Copilot on PC and Web, ChatGPT, Ollama	JAWS, NVDA
P11	Totally blind	35-44	United States	Vocational Rehabilitation Counselor	Siri, Alexa	Be My AI, Meta AI, Aira, Gemini	JAWS, NVDA, Narrator
P12	Legally blind	45-54	United States	Volunteer at Local Community	Alexa, Siri, Google Assistant	Be My AI, Picture Smart AI	JAWS
P13	Totally blind	55-64	Australia	Non-executive Board Member of a Advocacy Organization	Siri, Google Assistant	Be My AI	JAWS
P14	Legally blind	35-44	United States	Senior Director of Software Engineering	Cortana, Siri, Alexa	Seeing AI, Be My AI, ChatGPT, Picture Smart AI, Copilot on Web and Github	JAWS
P15	Totally blind	55-64	Australia	Lived Experience Educator	Google Assistant, Siri	Be My AI	JAWS, NVDA
P16	Legally blind	25-34	United States	Recruitment Coordinator	Siri, Alexa	ChatGPT, Copilot on Web	JAWS

Assistant, and Microsoft Cortana, and whether they were familiar with recently introduced GenAI tools such as ChatGPT, Gemini, and Copilot. Participants were asked to share their experiences using these tools. After assessing their awareness of these tools, we provided examples of the capabilities of GenAI (e.g., generating alt text for graphical information, summarizing and generating content/presentations, analyzing spreadsheets) and encouraged participants to think beyond the current capabilities of these tools.

We then presented an overview of survey findings, highlighting the themes that emerged as challenging areas (i.e., understanding content, accessing controls, authoring/editing, collaboration, and navigation) and the top three applications reporting the most issues (i.e., presentation tools, word processors and spreadsheets). Participants were asked to talk about the challenges they face, whether they consider GenAI to be beneficial in addressing these challenges, and to describe the capabilities they expect it to have or why they thought it might not be useful. We encouraged participants to think about how they would like to interact with these tools. We then inquired about the potential impact of GenAI on their productivity and independence. The interview concluded with questions about the potential concerns of using these AI tools. Participants were compensated with a \$50 Amazon gift card. The study was approved by our institutional ethics committee.

4.3 Data Analysis

The interviews were transcribed and analyzed using thematic analysis [23]. We employed an inductive approach with the first author open-coding all the transcripts with low-level descriptive codes. These low-level codes were then collectively refined with the remaining authors into high-level codes over seven weekly meetings. During these meetings, we discussed the associated excerpts till we reached consensus on the interpretation. The process involved adding, renaming, and removing codes. When a new code emerged,

the first author revisited all transcripts to identify relevant excerpts that supported the new code. Low-level codes such as “*generating alt text for images*” and “*tagging documents with proper headings*” were grouped under the high-level code “*transforming inaccessible content to accessible formats*”. In total, we had 101 low-level codes, that were grouped into 21 high-level codes. As a group, we then looked for relationships across the high-level codes and reflected on broader themes to describe these connections. After several iterations and discussions, we reached consensus on five overarching themes, which we elaborate on in the following section.

5 Findings

In this section, we present the five key themes that emerged from our analysis of the potential of GenAI assistance in productivity applications. Participants perceived GenAI as a valuable tool for enhancing workflow efficiency, increasing productivity, and supporting independence. Their preferences varied regarding the integration of GenAI into their workflows and interaction styles. Trust in GenAI emerged as a key concern, highlighting the need for verification methods for tasks performed by the AI. The following subsections elaborate on these themes, providing deeper insights into user needs and expectations.

5.1 Enhancing Workflow Efficiency and Productivity

A major concern for our participants was the additional time and effort required to complete productivity-related tasks when compared to their sighted peers. As P9 remarked, “*my goal is to be as productive as other people that I work with who are not blind. And it’s frustrating when I feel like I’m just as smart as they are, but it takes me longer to do things just because of the way things are set up*”. All participants anticipated that GenAI assistance, when accurate, would serve as a major time saver, reduce their workload, and

enable new opportunities in work environments. They valued its continuous and persistent availability, enabling them to complete tasks without waiting for sighted assistance.

Central to our findings was how participants were eager to maintain their existing workflows and use GenAI in a supportive manner to accomplish a wide range of tasks in productivity applications—particularly those that were inaccessible, time-consuming, or extremely challenging to perform with their SRs.

Understanding and applying visual formatting in documents: Most participants ($n=14$) noted that GenAI would be beneficial for formatting in documents, as they often struggled with overlapping elements, alignment, sizing, and styling issues. SRs either lacked information on these issues or were overly verbose when describing dimensions, causing participants to seek and wait for assistance from sighted colleagues. For example, P4 described a task involving adding a logo to a flyer and remarked, “AI could help me do that, and I don’t have to wait until someone comes over two days later to add it for me.” Participants also expressed a desire for GenAI to have the capability to identify and correct formatting errors or provide guidance on how to fix them. P13 expressed, “I think having an assistant that can explain, this looks like this and this looks like that, or just edit it so that, all the headings are in the right side and there’s the right spaces in your document [...] something almost like an editorial assistant in a sense, if it could understand that sort of level of detail and either do it or put you in the place to do it or give you the commands to do it.”

Some participants ($n=9$) desired assistance during the document creation process to specify layout preferences. P14 expressed their preference when creating presentations, “it would be great if there was a way to use prompts and tell it, I want a bulleted list on the right, or, I need, like four boxes on the screen and I need arrows pointing from the box on the upper left to the box in the lower right”. P16 shared concerns about being judged for visual misalignments by colleagues. P9 suggested GenAI could also help by recommending visual improvements, like breaking up text-heavy slides.

Understanding and generating graphics: The ability to describe graphical information was another feature that participants ($n=10$) found to be highly beneficial. While some participants already used tools such as Picture Smart AI and Be My AI, they appreciated these tools’ capability to not only generate descriptions but also go beyond them to provide further information and clarifications through follow-up questions. P4 shared, “I would say, ‘what is this an image of?’ And if it says it’s a chart of produce sales in different states for example, then ‘what are the different categories?’ and then, if I knew it was broccoli, carrots [...], and it was Arizona, California [...] then I could say, ‘what’s the relevant data for California?’”. Controlling response verbosity was deemed useful ($n=3$). P10 explained that image descriptions are sometimes overly detailed and preferred control over their length, stating, “I don’t want a paragraph; I just want a quick 10-word description’ [...] There’d be, short descriptions, medium descriptions, long descriptions”. Additionally, three participants mentioned that they look forward to having a feature that could describe videos (e.g., lectures, YouTube videos) due to lack of audio descriptions. P7 said, “I would really like to have a mode where it continuously takes screenshots and describes changes in the

screen, like just new things that appear on the screen or happen on the screen”.

Participants also expressed interest in generating their own graphical content. They valued the ability to generate and insert images based on content or prompts, as well as the capability to describe and ask questions about the generated images. P6 highlighted that this feature could enable them to design book covers for their written works, which currently acted as a barrier to publishing, describing it as a *liberating* capability. P13 noted that this would enable them to add images to their presentations—something they had not typically considered doing before.

Understanding and extracting information from large documents: Participants ($n=11$) believed that GenAI could assist in extracting information from large documents. They expressed interest in obtaining a summary of the document’s structure and content. P15 said, “I think it would be really valuable in a spreadsheet where it could say, look, this is a 10 column spreadsheet with the horizontal column headers blah blah [...]. And down the Y axis, it shows dates”. Participants ($n=5$) also valued the ability to ask questions about specific content within the document. For instance, P11 said, “For a sales document, I may be asking like, ‘What were the monthly profits? What were the monthly expenditures?’” Participants ($n=8$) preferred asking questions about totals, minimums, maximums, averages, and trends in spreadsheets rather than using data analysis features like formulas or sorting, due to cell-by-cell navigation of the SR and the multiple steps required to access these features. They also highlighted the value of being able to request navigation to relevant sections, due to the inefficiencies of navigating with the SR. P9 remarked, “so say it had a big document, it summarized it for me, and there’s a topic that talks about these three things. If I could be like, ‘take me to the part where it talks about whatever the second one is,’ and then it would take me there and I could actually read the real document.”

Understanding and tracking changes made by others: Participants ($n=7$) highlighted the potential of GenAI for improving document editing and collaboration, particularly tracking changes made by others. They found it challenging to differentiate between the original content and the suggested edits. P9 noted, “Because a lot of times, it’ll be like, something was deleted, and something was added. And it’s just a whole bunch of words. [...] I just want to hear the before and after, so I can decide, is this better, or is what was already there better.” Three participants recommended receiving a summary of changes made by others while they were not on the document to avoid missing important edits and to reduce the time and effort required to find them. P14 highlighted the complexity of navigating comments, saying it involves “a lot of keyboard commands and maneuvering”. They suggested an assistant that could provide context to the comment, read it aloud, and even help with replying. P2 recommended that the assistant could notify them of any unresolved comments in documents.

Creating tables and writing formulas in spreadsheets: Participants ($n=3$) believed that GenAI could assist them in creating tables more efficiently by using natural language commands, enabling them to specify the column titles and their formats. P15 stated, “It would be really good if I could say to it, ‘set up a spreadsheet with

these columns. This one's a numeric column, and that one's a text column, and that one's a date column.” Some participants ($n=4$) also mentioned that they could create formulas by mentioning the operation and the columns. P10 remarked, “add all of what's in column A and put the total in column B. I don't remember the formula. Can you write it out for me?”

In addition to GenAI supporting application-specific capabilities, they envisioned it working across multiple applications. For example, P6, who had lost their job due to inaccessible software in the workplace, remarked,

“With AI, I could take my spreadsheet and create a PowerPoint presentation out of Excel—let's say a pie chart—have it describe the segments to me of the pie chart, help work with me to title it, insert it into a PowerPoint presentation that I'm creating, work on a logical flow. I could ask questions along the way that would allow it to pull other data [...] I mean, it's just, the sky's the limit with this.”

Participants ($n=5$) emphasized that GenAI tools can offer greater benefits to them than to sighted users. They highlighted that such tools can enable interaction with inaccessible software, reduce the additional effort and time spent learning keystrokes—which often vary across applications—and mitigate the disadvantages associated with the inability to visually interact with applications and their content. For instance P16 stated,

“I'm excited to see the integrations with productivity, specifically for disabled individuals [...] I think that, especially for someone that has a lot of disadvantages because of disabilities, it adds a couple more advantages back for me, more so than other people. Other people might be able to just visually scan that document—no problem, whereas for me, an AI that can scan and summarize that document really quick for me could be a game changer.”

5.2 Supporting Independence

All participants stated that they would prefer seeking assistance from GenAI rather than asking people, as they valued their independence and disliked burdening others or wasting their time. They appreciated its ability to engage in back-and-forth conversations ($n=4$), answer unlimited questions without frustration ($n=4$), and provide the freedom to explore without judgment ($n=2$). Most participants ($n=8$) commented that completing tasks faster was a key priority, with no concern over whether they or the AI performed the task. Some ($n=3$) even found it more satisfying to achieve tasks with the help of AI assistance. P14 remarked, “I would find that very satisfying if I was using AI, I could care less if I created it myself or I used AI. It's still me manipulating AI. It's just another tool.”

Beyond performing tasks and extracting information (as highlighted in Section 5.1), participants noted that GenAI could enhance their independence by improving the effectiveness of SR use:

Learning to perform tasks with the help of AI: Participants ($n=13$) discussed the importance of learning to perform tasks independently with their SRs and viewed GenAI as a tool to support this by providing instructions and keyboard shortcuts. P15 suggested,

“You could actually have an AI system to teach you, like to give you lessons to learn it or even the option to say, ‘Hey, [...] how do I do the formula to add up a column?’” Participants noted that the choice of using the AI to perform a task or to teach how to complete the task with a SR depended on factors such as task urgency, frequency, complexity, and the user's expertise. For example, P10 said, “I would first ask the AI, ‘How would you do this?’ because that's what I do on Google all the time. Then, I would get some how-tos and read those. But if I were in a hurry and knew the AI could do it, I would just tell it to do it.” However, P2 and P8 indicated that they would always prefer learning tasks with the help of AI rather than asking it to perform them. P8 remarked, “I like to maintain my independence, and that means over time, I'm less reliant on AI because I've done it and learned it, as opposed to it automatically doing it.” P2 mentioned they are already using ChatGPT to learn about steps and keystrokes for various tasks. They also emphasized that learning to perform tasks independently remained important in case GenAI became unavailable due to technical issues, such as internet connectivity problems.

Transforming inaccessible content into accessible formats:

Participants suggested that GenAI could make documents more accessible and therefore support independent exploration with a SR. Eight participants noted that they frequently received digital documents (e.g., word processor files, spreadsheets, and presentations) not designed with accessibility in mind and identified the potential for GenAI to convert these into accessible formats that could be independently navigated using their SRs. For instance, P16 mentioned frequently encountering résumés without properly tagged headings and describing how, when using their SR, “it's just a giant chunk of text.” They proposed that GenAI could add appropriate headings to such documents, allowing them to skip sections and navigate directly to the information they need. Participants wanted the ability to access data contained in images through their SRs. Apart from generating alternative text, some participants ($n=2$) proposed that GenAI could extract data from images of tables or charts and convert it into spreadsheet tables. P13 preferred the ability to navigate chart elements (e.g., axis labels, titles) and data points over asking questions about the data. They explained, “So, rather than it just saying, ‘The top performer this month was this [...]’ I would probably want to actually access the figures to make that judgment myself.” They described the assistance as a “secret assistant” that could help them complete tasks without requiring everything to be fully accessible. However, they also acknowledged that relying on such capabilities might discourage the creation of accessible documents, which could negatively impact users who lack access to or cannot afford access to these tools.

While all participants were open to using GenAI, some were judicious about the areas where they sought assistance. For instance, P5 and P16 noted that they would not use AI for tasks they were emotionally invested in such as writing. They preferred completing these tasks themselves, stating that it “wouldn't feel genuine” if they were credited for AI-generated output. Additionally, several participants ($n=4$) expressed less satisfaction if AI handled tasks they could perform independently using their SR. P8 likened it to “giving someone else your work to do.” Some participants ($n=4$) also highlighted that they did not want to rely too heavily on AI;

however, most of them commented that relying on AI to some extent was acceptable, as it would enhance their quality of life by saving time.

5.3 Integrating GenAI

To accommodate their workflows (as described in Sections 5.1 and 5.2), all participants highlighted the importance of GenAI working seamlessly with their SRs. They noted that SRs still play a crucial role for two main reasons: (1) SRs are essential for input, content access, and verifying output ($n=13$); and (2) some tasks are faster to perform using SR shortcuts rather than commanding and waiting for the responses ($n=6$). Participants envisioned GenAI integration into their workflows in various ways:

Integrated into the SR: Participants ($n=5$) preferred this method due to their familiarity with SRs and the control SRs provide, as well as their aversion to interacting with multiple programs. They valued the ability to access assistance across various applications, consistency of key commands, and the seamless collaboration between the assistant and the SR. P14 highlighted that this integration would support better communication and help mitigate the issue of “*the screen reader and the assistant talking over each other.*” Additionally, participants believed that integrating assistance within SRs could be designed for non-visual access, making it more accessible and usable.

Integrated into the system independent of the SR: Several participants ($n=6$) preferred having a consistent interface and the ability to access any document or application across the operating system (OS). For instance, P7 noted that such integration would mitigate the need to transfer data between application-specific assistants, “*let’s say you write an outline for a presentation in Microsoft Word and you want that to go in PowerPoint. You would have to copy that and paste it into PowerPoint’s Assistant. So it would just be better to have everything in one OS assistant.*” Furthermore, participants highlighted that a universal assistant could learn from information across various documents and past interactions. For example, P16 noted that this would eliminate the need to repeatedly explain how they want their documents formatted. Four participants highlighted that, since the assistant and SR are not interdependent, this separation would prevent GenAI failures from disrupting the SR. P4 said, “*I wouldn’t want it to be something I had to get updates for, that could cause any more hiccups in the SR than there already are.*” Additionally, two participants noted that independent assistance would be accessible with multiple types of SRs.

Integrated into the productivity application: Some participants ($n=3$) noted that assistant integrated within an application would perform better and possess more knowledge about specific application features. For instance, P4 believed that the assistant would be more aware of different application versions and updates. Additionally, P3 noted that an application-level assistant would receive updates more frequently than OS-level assistants.

While most participants preferred one integration method, they also saw benefits in others. One participant preferred having multiple methods, while another felt that the method was less important as long as the assistance could be activated on demand.

5.4 Interaction Styles and Preferences

Almost all participants ($n=14$) preferred having both voice and text input options for giving commands to GenAI. Participants valued the ability to interact with it without disrupting their focus on the current task ($n=5$), noting that voice input was more convenient in such instances. P4 remarked, “*if I had to type to this assistant, I’d probably have to be going out of the application to some edit box to respond to it and then find my place back in the document.*” However, participants stated that in cases where command accuracy was crucial, they would prefer typing. This was attributed to concerns that accents and speech styles could be misinterpreted by such tools. Participants ($n=12$) also noted that the choice of input method depends on their setting, whether in a shared work environment or at home.

Participants expressed varied preferences for GenAI speech styles. Of the participants, nine preferred a machine-like (i.e., robotic) voice for the assistance, with some favoring a machine voice distinct from their SRs. Six participants opted for a more human-like voice, and one participant preferred having both options. While some found the human-like voice more understandable and a helpful way to distinguish between SR and assistant outputs, others felt it would reduce efficiency. P2 explained that human-like voices adds more latency with pauses, making them feel *sluggish* compared to machine-like voices commonly used in SRs. Participants ($n=6$) needed to adjust the assistant’s speech rate based on the context and urgency of the task. For instance, P4, who preferred a human-like voice, stated they would increase the speech rate to be more productive at work. In addition to accessing GenAI output through speech, participants ($n=12$) also wanted the content (e.g., summaries, answers to questions) to be recorded in text, allowing them to review the output with their SRs. P4 noted that it would allow them to check details such as spelling, character by character, which they stated was not possible with a “*voice hanging in the air.*”

Participants expressed contrasting views on the ideal personality for their GenAI assistant. P12 and P16 preferred it to have a friendlier tone and a personality similar to a colleague. P16 elaborated on their preference for anthropomorphizing the assistant, “*I do like ones that are a little bit more human, like they do small talk with you, where it’s like, ‘how’s your day going? [...] what can I help you with today?’*” P6 mentioned that the assistant’s tone holds credibility and should be customizable—allowing users to choose between a more technical or conversational style. For example, when creating a book cover, they preferred the assistant to act like a graphic designer collaborating with them to get the job done. However, six participants strongly disliked the assistant having informal personalities, emphasizing that it is a tool meant to deliver specific information and complete the tasks requested. P14 highlighted that an assistant attempting to be more human-like would add extra verbosity, negatively affecting their productivity.

Participants preferred having multiple methods for activating the assistant to give commands. Some ($n=12$) favored a keyboard shortcut or dedicated button for activation, while others ($n=3$) found a wake word more convenient. Four participants preferred to have the assistant running in the background, ready to be triggered when needed. For example, P16 wanted it to function “*kind of like the Amazon device, where it’s kind of in a standby, but it can be ready to*

go at any time.” While most participants preferred explicit activation, two favored continuous activation throughout an interaction for certain tasks (e.g., writing a story). However, P14 disliked continuous activation, noting that unintended voice interactions or SR output could be misinterpreted as input by the assistant.

5.5 Trust in GenAI

All participants raised concerns about the inaccuracies and hallucinations of GenAI. They shared instances where GenAI provided false information. For example, P1 said, “I took a photo of the husky that we look after. And I said, ‘what’s this picture of?’ It’s a cat lying on the bed. And I went back saying, ‘are you sure it’s not a dog?’ And they said ‘no, by the shape of the ears, shape of the face, texture of the fur, it’s a cat.” Due to such confident responses given by GenAI, participants found it hard to fully trust these tools. P15 was hesitant to use certain features due to similar stories from other blind users. However, they also shared positive experiences where GenAI performed well, increasing their trust. P9, for example, praised JAWS Picture Smart AI for accurately describing an infographic in a presentation, which gave them a better understanding when presenting to others. Participants ($n=9$) felt that GenAI tools were more accurate for certain tasks, noting fewer errors in tasks based on already provided content, such as summarizing documents, compared to tasks that required interpretation, like image descriptions.

Participants ($n=6$) perceived GenAI tools as having negative biases, which led to a lack of trust in its ability to provide a representative answer. For instance, P1 recounted hearing from another blind person who asked GenAI tools, “Do blind people make good parents?” and received a very negative response. They expressed concern that GenAI’s answers are shaped by widespread biases and discrimination found online, stating, “I’m worried about the whole garbage in, garbage out stuff.” Participants ($n=4$) stressed the importance of critical thinking when working with GenAI. P7 also noted that these models are too general and may lack detailed knowledge about blind people, accessibility, or the technologies they use, stating, “if you try to make it write a lesson plan for a screen reader, it may not know very much at all about screen reader commands”.

Although participants expressed some trust in GenAI output, they emphasized the importance of having the ability to verify it. Most participants ($n=12$) preferred being directed to the original content to verify text-based output. P7 and P14 suggested that providing confidence levels for image-based outputs could help gauge accuracy. For example, information such as “there’s a cat in this picture, I’m 60% sure of it” was considered desirable. Participants described various methods they currently use to verify accuracy, including running the same command through multiple models (e.g., cross-checking image descriptions generated by ChatGPT and Claude in Picture Smart AI) ($n=5$), repeating the command within the same model ($n=1$), magnifying the content ($n=3$), and relying on their prior knowledge of the content ($n=4$). For verifying tasks performed by GenAI, participants ($n=9$) preferred receiving detailed verbal feedback upon completion. P15 provided an example involving sorting a column in a spreadsheet, stating, “I’d want it to say the name of the column that has been sorted. The first cell in the column is Alfred and the last cell in the column is Ziggy, just so I knew

that it had actually done it properly.” Participants also noted that in cases requiring 100% accuracy or involving sensitive information, they would seek human verification. However, in situations where verification methods were unavailable, P16 stated they would “just have to take a leap of faith” and trust the output to be accurate.

Similar to prior studies on use of GenAI [6], privacy concerns related to sharing confidential information were echoed by all of our participants. Most participants ($n=12$) lacked trust in GenAI tools providers’ and their data protection policies, primarily due to past reports of data breach incidents. Participants ($n=12$) mentioned avoiding these tools for tasks involving personal or confidential information, fearing their data could be leaked or misused. A few participants expressed less concern about the providers themselves but were more worried about third-party hacking. P7 and P14 suggested using local models for handling sensitive data. Conversely, two participants were less worried, feeling that much of their personal data was already exposed in various ways. Participants recommended increased transparency from AI providers on data ownership, source code, and training data ($n=5$), clearer and more concise policies ($n=5$), and options for users to access and delete interaction history ($n=3$). However, some participants ($n=4$) noted that privacy concerns might lead workplaces to restrict GenAI use, thereby limiting its benefits for them.

Overall, despite the privacy risks and considerations around trust, all participants indicated that they would still choose to use GenAI, as they believed the benefits outweighed the concerns. They were optimistic about it being a game-changer for them. P12 remarked, “AI is going to change the way we do stuff as blind people. It’s the best thing that ever came out for us.”

6 Discussion and Future Opportunities

6.1 The Potential Usefulness of GenAI

Overwhelmingly, our participants were eager to use GenAI in productivity applications despite concerns around trust, verification and privacy. They not only valued the independence it afforded but also the effort and time savings when performing tasks. Although a few participants were worried about over-reliance on these tools, they acknowledged the ability to carry out tasks that were previously impossible or extremely difficult to perform using their SRs. Overall, GenAI was regarded as an enabler, particularly in workplace settings, that would help reduce the disability divide—the gap in access to opportunities and resources compared to sighted individuals. Notably, some participants were worried that workplaces might limit the use of these tools due to privacy and security concerns.

Participants’ ($n=3$) initial perceptions of GenAI’s usefulness increased between the survey and the interview, likely due to the widespread adoption and increased awareness of these tools in the intervening months. The interview participants were optimistic that GenAI could help with the high level challenges identified in the survey (Sections 3.3.2 and 5.1), namely:

- (1) *Understanding content*: Describing graphical data and document formatting, answering questions about data, summarizing content, and transforming inaccessible content into accessible formats.

- (2) *Accessing controls*: Providing keystrokes for different tasks and providing the ability to use natural language commands to access features, eliminating the need to navigate through multiple menu items.
- (3) *Authoring and editing*: Generating images, creating tables and formulas, and formatting documents and suggesting improvements.
- (4) *Navigation*: Providing the ability to jump to different parts of the document using natural language commands.
- (5) *Collaboration*: Providing summaries of changes made by others in shared documents and assisting with understanding edits and comments from collaborators.

Addressing these challenges with GenAI benefits not only SR users but also sighted individuals. However, a unique need for SR users is the ability to transform inaccessible content into accessible formats (Section 5.2). While our participants recognized the usefulness of this feature, they also emphasized that having GenAI does not eliminate the need for content to be born accessible. This is due to several factors, such as limited access to or affordability of GenAI tools, as well as a lack of trust caused by inaccuracies and hallucinations in AI-generated outputs [6, 9]. However, GenAI could help in this process, particularly in the document creation workflow to assist authors in tagging and formatting accessible content. For example, GenAI functionalities such as identifying heading levels from plain text [104, 105], transforming inaccessible table and chart data into SR-accessible formats [52, 72], and generating alt text for images [33] could be integrated into productivity tools or implemented as suggested prompts to guide authors in creating accessible documents. Authors could then review and verify the document’s accessibility before sharing it with blind users.

6.2 GenAI’s Role in Accommodating Existing Workflows

Regardless of the interaction methods presented by GenAI (e.g., voice or text-based chat), all of our participants wanted these tools to be supportive of their existing SR workflows (Section 5.3). This echoes findings of Vtyurina et al. [115] and Phutane et al. [93], who also emphasize the continued importance of SRs, even when conversational assistance is available. Our participants valued the control that SRs provided, their familiarity with using them, and highlighted that for some tasks, it was more efficient to use SRs than to command GenAI and wait for an output.

SR users wanted to make decisions about what the GenAI should do but were also open to suggestions from it. They perceived GenAI as a supplementary helper in this mixed-initiative setting [10], where humans or computers contribute what they are best suited for at the most appropriate time. GenAI in this case would be capable of supporting them in their existing workflows by always being present in the background or through explicit activation using wake words or keyboard commands. We identified four main roles SR users expected GenAI to play when supporting their workflows. This includes:

- *Information provider*: Extracts information from documents and provides answers to questions based on their content, such as generating a summary upon request.

- *Task performer*: Performs the requested tasks on behalf of the user, such as sorting a column in a table, formatting a document according to user requirements, or transforming a document into accessible formats.
- *User guide*: Provides users with requested instructions or keyboard commands on how to perform tasks using their SR. For example, this could include providing a step-by-step guide to creating a formula in a spreadsheet using keyboard commands.
- *Reviewer*: Identifies mistakes or provides improvements for tasks performed by SR users. This could include informing users about formatting errors or proofreading a document and suggesting fixes or improvements.

These roles can be mixed and combined in the same workflow. How these roles are used are dependent on several factors: the urgency of the task, the complexity of performing it with the SR, the user’s knowledge of the task, and the required level of accuracy. When users are in a hurry to complete a task, they may ask GenAI to perform it rather than learning the steps themselves. However, when accuracy and correctness are at stake, users may use the *reviewer role* to identify any mistakes or the *user guide* role to receive instructions for doing the task themselves.

Our participants also highlighted that, when using GenAI assistance, they preferred not to shift their SR focus from the current task. Shifting the SR focus to another screen or text field and typing a command was perceived as an interruption to their workflow, making it difficult to return to their original position. They found voice commands more convenient for interacting with AI [61, 115]. However, there were exceptions: in shared environments, participants preferred not to use voice commands, or when they required commands to be 100% accurate with no room for misinterpretation, they still preferred text input. This suggests that transitions between voice and text based interactions with GenAI need to be seamless in order to support their current workflows.

6.3 Supporting Trust and Verification

The introduction of GenAI has brought about a paradigm shift from a command-based to an intent-based interaction paradigm, where users tell the computer “*what to do*” rather than “*how to do it*” [85]. While this shift helps both sighted and blind people achieve tasks more efficiently, blind people benefit significantly as it reduces frustrations with extensive maneuvering and keyboard commands. However, in this type of interaction, as the locus of control is entirely assigned to the GenAI when providing information or carrying out a task, it may be prone to inaccuracies and hallucination that users might find difficult to identify and correct [85]. While all users—blind and sighted alike—need to grapple with these issues, they are exacerbated for SR users (Section 5.5). In some instances, sighted users will be able to visually identify issues and verify the accuracy of the information or the task, but doing so with a SR can be impossible or significantly slower and more challenging. This is an area that requires further research, but potential methods for building trust and verification for SR users could be accomplished through:

- *Providing contextual feedback on tasks:* Similar to prior work highlighting the importance of notifications upon completing response generation in GenAI tools [6], our participants emphasized the need for detailed feedback on task completion to ensure accuracy. For example, when sorting a column in a spreadsheet, participants suggested being informed about the sorted column, the sort order, and the first and last entries.
- *Communicating confidence levels and uncertainty:* Our study, along with prior work [9, 74], highlights that explainable AI (XAI) features such as confidence ratings can assist blind users in mitigating over-trusting AI technologies. While previous research has primarily focused on tasks related to understanding images, future work could explore the benefits of these methods in productivity-related tasks. Providing confidence levels alongside task feedback could enable more efficient decision-making regarding verification. However, users may also wish to understand how accuracy ratings are computed [9]. Additionally, many participants found it useful to have the ability to direct their SR focus to the original source of the content for verifying the accuracy of the GenAI-generated responses.
- *Cross-checking accuracy between models:* Our participants suggested cross-checking the accuracy of image descriptions using multiple models. While this approach has been shown to be useful for verifying image-based information [6, 9], its effectiveness in text-based or task-oriented scenarios requires further exploration.

6.4 Integration and Personalization of GenAI

GenAI is increasingly being integrated into various applications, leading to a proliferation of AI assistants that function differently. While this lack of standardization poses usability challenges for all users in adapting to multiple assistants [32], it is particularly problematic for SR users. Assistants with varying keystrokes and interfaces impose additional cognitive load on SR users, making interactions more complex. This challenge is reflected in our findings (Section 5.3), where the majority of participants expressed a preference for a more universal approach to accessing GenAI assistance. Addressing this issue requires not only better interoperability standards between different systems but also consistent interaction methods that seamlessly accommodate and support SR use.

Regardless of how GenAI is integrated, our study aligns with prior work emphasizing the importance of allowing blind users extensive customization of AI interaction styles [31, 93, 96]. For most of our participants, productivity gains were the top priority, leading them to prefer a more machine-like voice and dislike of AI with informal personalities (Section 5.4). However, some participants also expressed a preference for personalities resembling a colleague or a specialist (e.g., a graphic designer), noting that such personas could create a sense of credibility. While this could engender confidence in the system, it could also potentially result in over-trust and increased reliance on AI, as highlighted by Adnin and Das [6] and Liao and Vaughan [70].

6.5 Including Blind Perspectives in GenAI

Our analysis aligns with previous studies that have highlighted harmful disability representations in content produced by GenAI tools [6, 9, 43, 45, 46]. For instance, receiving negatively biased responses to questions such as “*Do blind people make good parents?*” serves as a clear example of ableist biases against blind people (Section 5.5). Echoing findings in prior work [6, 43], our participants also emphasized that this is a result of the ‘garbage in, garbage out’ phenomenon, stemming from the bias and discrimination prevalent in the training data. Recent research has proposed ways to mitigate these biases, including evaluating models [76, 119], implementing guardrails [103], and creating more representative datasets [11, 18, 59, 113], and conducting AI audits in collaboration with blind people [3, 43, 107].

Our study also highlights the need to train models to better understand how blind people use assistive technologies, such as SRs. For instance, one of our participants emphasized the importance of providing user guides and tutorials specifically designed for non-visual access, detailing steps to perform tasks using keyboard interactions. Another use case identified in our study, which is also found in prior work [46, 62, 95], is the conversion of inaccessible document content into SR-accessible formats. However, informal testing suggests that current models are ineffective in performing this task. Training models with both properly formatted and tagged accessible documents, as well as inaccessible documents with explicit annotations, could be a potential approach. However, further research is needed to explore methods for teaching models to identify accessibility issues and generate appropriate tags and alt text that align with the needs of SR users.

7 Limitations

Our participants were highly educated and early adopters of current technologies. Most were from English-speaking countries, such as the U.S. and Australia, and therefore, our findings may not fully reflect the perspectives of the broader blind population. As our survey was conducted during the early months of GenAI’s public availability, participants likely had limited knowledge of its capabilities or how to use such tools to address their specific needs. Given the rapid evolution of GenAI capabilities, our research captures a particular snapshot in time of people’s attitudes. Additionally, our participants were imagining how they might use GenAI with productivity applications and how they would prefer to use it; their opinions might change when interacting with an actual system. Lastly, we acknowledge our positionality as sighted accessibility researchers, which we recognize as a limitation, particularly in a study focused on the experiences and perspectives of blind users regarding AI technologies.

8 Conclusion

We conducted a survey with 99 SR users to gain insights into their use of productivity applications, focusing on obtaining a holistic understanding of the challenges they face. In line with previous studies, we found that users experienced difficulties accessing content and controls, authoring and editing, navigating interfaces, and collaborating. These challenges significantly impacted their productivity, independence, and emotional well-being. Our survey and

interviews with 16 SR users revealed that the SR community held a positive outlook on the potential of GenAI assistance in productivity applications, believing it could address many of the challenges they currently encounter. Participants viewed GenAI as an adjunct to their existing workflows, rather than a replacement for SRs, and emphasized the importance of supporting verification. The findings presented in this paper offer valuable insights for technology providers, helping them improve the experience for SR users. Enhancing application accessibility and usability will empower SR users to independently access information, ultimately improving educational and employment opportunities and their overall quality of life.

Acknowledgments

We thank our participants for their valuable contributions in sharing their experiences and insights. We also thank Dr. Leona Holloway for her assistance with participant recruitment, Kate Stephens for evaluating the screen reader accessibility of the survey, and the members of the Monash Inclusive Technologies group for their support and valuable discussions.

References

- [1] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. 2018. "Siri Talks at You" An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 249–258.
- [2] Ali Abdolrahmani, Kevin M Storer, Antony Rishin Mukkath Roy, Ravi Kuber, and Stacy M Branham. 2020. Blind leading the sighted: drawing design insights from blind users towards more productivity-oriented voice interfaces. *ACM Transactions on Accessible Computing (TACCESS)* 12, 4 (2020), 1–35.
- [3] Ariana Aboulafia and Bonnielin Bogen, Miranda Swenorbe. 2024. To Reduce Disability Bias in Technology, Start With Disability Data. Retrieved December 5, 2024 from <https://cdt.org/wp-content/uploads/2024/07/2024-07-23-Data-Disability-report-final.pdf>
- [4] NV Access. 2024. NVDA. Retrieved March 22, 2024 from <https://nvda.groups.io/g/nvda>
- [5] NV Access. 2025. NVDA. <https://www.nvaccess.org>
- [6] Rudaiba Adnin and Maitraye Das. 2024. "I look at it as the king of knowledge": How Blind People Use and Understand Generative AI Tools. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (St. John's, NL, Canada) (ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 64, 14 pages. doi:10.1145/3663548.3675631
- [7] Aira. 2025. Aira. <https://aira.io/>
- [8] Eman A Alasadi and Carlos R Baiz. 2023. Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education* 100, 8 (2023), 2965–2971.
- [9] Rahaf Alharbi, Pa Lor, Jaylin Herskovitz, Sarita Schoenebeck, and Robin N. Brewer. 2024. Misfitting With AI: How Blind People Verify and Contest AI Errors. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility (St. John's, NL, Canada) (ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 61, 17 pages. doi:10.1145/3663548.3675659
- [10] James E Allen, Curry I Guinn, and Eric Horvitz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14, 5 (1999), 14–23.
- [11] Bedour Alshaigy and Virginia Grande. 2024. Forgotten Again: Addressing Accessibility Challenges of Generative AI Tools for People with Disabilities. In *Adjunct Proceedings of the 2024 Nordic Conference on Human-Computer Interaction*. Association for Computing Machinery, New York, NY, USA, 1–6.
- [12] Amazon. 2024. Alexa. Retrieved January 10, 2024 from <https://developer.amazon.com/alexa>
- [13] Apple. 2024. iWork. Retrieved January 10, 2024 from <https://www.apple.com/au/iwork/>
- [14] Apple. 2024. Siri. Retrieved January 10, 2024 from <https://www.apple.com/siri/>
- [15] Sparkling Apps. 2025. PiccyBot. <https://www.piccybot.com/>
- [16] Blind Citizens Australia. 2024. Blind Citizens Australia. Retrieved March 22, 2024 from <https://www.bca.org.au/>
- [17] Matthew J Baker, EM Nightingale, and Suzy Bills. 2021. An editing process for blind or visually impaired editors. *IEEE Transactions on Professional Communication* 64, 3 (2021), 275–287.
- [18] be my eyes. 2024. Be My Eyes Announces New Collaboration with Microsoft to Help Train Inclusive AI Models. Retrieved December 4, 2024 from <https://www.bemyeyes.com/blog/be-my-eyes-announces-new-collaboration-with-microsoft>
- [19] be my eyes. 2024. Be My Eyes for Windows. Retrieved September 12, 2024 from <https://www.bemyeyes.com/be-my-eyes-for-windows>
- [20] be my eyes. 2024. Introducing: Be my AI. Retrieved September 12, 2024 from <https://www.bemyeyes.com/blog/introducing-be-my-ai>
- [21] Stacy M Branham and Shaun K Kane. 2015. The invisible work of accessibility: how blind employees manage accessibility in mixed-ability workplaces. In *Proceedings of the 17th international acm sigaccess conference on computers & accessibility*. 163–171.
- [22] Stacy M Branham and Antony Rishin Mukkath Roy. 2019. Reading between the guidelines: How commercial voice assistant guidelines hinder accessibility for blind users. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 446–458.
- [23] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association.
- [24] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33 (2020), 1877–1901.
- [25] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, Giulio Mori, and Victor MR Penichet. 2010. Accessing Google docs via screen reader. In *Computers Helping People with Special Needs: 12th International Conference, ICCHP 2010, Vienna, Austria, July 14-16, 2010. Proceedings 12*. Springer, 92–99.
- [26] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, Giulio Mori, and Victor MR Penichet. 2014. Collaborative editing: collaboration, awareness and accessibility issues for the blind. In *On the Move to Meaningful Internet Systems: OTM 2014 Workshops: Confederated International Workshops: OTM Academy, OTM Industry Case Studies Program, C&TC, EI2N, INBAST, ISDE, META4eS, MSC and OnToContent 2014, Amantea, Italy, October 27-31, 2014. Proceedings*. Springer, 567–573.
- [27] Alexia Cambon, Brent Hecht, Ben Edelman, Donald Ngwe, Sonia Jaffe, Amy Heger, Mihaela Vorvoreanu, Sida Peng, Jake Hofman, Alex Farach, Margarita Bermejo-Cano, Eric Knudsen, James Bono, Hardik Sanghavi, Sofia Spatharioti, David Rothschild, Daniel G. Goldstein, Eirini Kalliamvakou, Peter Cihon, Mert Demirel, Michael Schwarz, and Jaime Teevan. 2023. *Early LLM-based Tools for Enterprise Information Workers Likely Provide Meaningful Boosts to Productivity*. Technical Report. Microsoft. <https://www.microsoft.com/en-us/research/publication/early-llm-based-tools-for-enterprise-information-workers-likely-provide-meaningful-boosts-to-productivity/>
- [28] Yoonha Cha, Isabela Figueira, Jessy Ayala, Emory James Edwards, Joshua Garcia, André van der Hoek, and Stacy Marie Branham. 2024. "Do You Want Me to Participate or Not?": Investigating the Accessibility of Software Development Meetings for Blind and Low Vision Professionals. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–17.
- [29] Yoonha Cha, Victoria Jackson, Isabela Figueira, Stacy Marie Branham, and André Van der Hoek. 2024. Understanding the Career Mobility of Blind and Low Vision Software Professionals. In *Proceedings of the 2024 IEEE/ACM 17th International Conference on Cooperative and Human Aspects of Software Engineering*. 170–181.
- [30] Yuyan Chen, Qiang Fu, Yichen Yuan, Zhihao Wen, Ge Fan, Dayiheng Liu, Dongmei Zhang, Zhixu Li, and Yanghua Xiao. 2023. Hallucination detection: Robustly discerning reliable answers in large language models. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. 245–255.
- [31] Dasom Choi, Daehyun Kwak, Minji Cho, and Sangsu Lee. 2020. "Nobody speaks that fast!" An empirical study of speech rate in conversational agents for people with vision impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [32] Christopher Clarke, Karthik Krishnamurthy, Walter Talamonti, Yiping Kang, Lingjia Tang, and Jason Mars. 2024. One Agent Too Many: User Perspectives on Approaches to Multi-agent Conversational AI. *arXiv preprint arXiv:2401.07123* (2024).
- [33] Maitraye Das, Alexander J Fiannaca, Meredith Ringel Morris, Shaun K Kane, and Cynthia L Bennett. 2024. From Provenance to Aberrations: Image Creator and Screen Reader User Perspectives on Alt Text for AI-Generated Images. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–21.
- [34] 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 (2019), 1–26.

- [35] Maitraye Das, Thomas Barlow McHugh, Anne Marie Piper, and Darren Gergle. 2022. Co11ab: Augmenting accessibility in synchronous collaborative writing for people with vision impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–18.
- [36] 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* 29, 2 (2022), 1–42.
- [37] Iyad Abu Doush and Enrico Pontelli. 2010. Non-visual navigation of tables in spreadsheets. In *Proc. of ICCHP*. 108–115.
- [38] Iyad Abu Doush and Enrico Pontelli. 2013. Non-visual navigation of spreadsheets: Enhancing accessibility of Microsoft Excel™. *Universal access in the information society* 12 (2013), 143–159.
- [39] Christof Ebert and Panos Louridas. 2023. Generative AI for software practitioners. *IEEE Software* 40, 4 (2023), 30–38.
- [40] Benjamin G Edelman, Donald Ngwe, and Sida Peng. 2023. Measuring the Impact of AI on Information Worker Productivity. Available at SSRN 4648686 (2023). <https://ssrn.com/abstract=4648686>
- [41] D Gareth Evans, Tim Diggle, Sri H Kurniawan, and Paul Blenkorn. 2003. An investigation into formatting and layout errors produced by blind word-processor users and an evaluation of prototype error prevention and correction techniques. *IEEE Transactions on neural systems and rehabilitation engineering* 11, 3 (2003), 257–268.
- [42] FreeLists. 2024. vip-l: Vision Impaired People's List. Retrieved March 22, 2024 from <https://www.freelists.org/list/vip-l>
- [43] Viniitha Gadiraju, Shaun Kane, Sunipa Dev, Alex Taylor, Ding Wang, Emily Denton, and Robin Brewer. 2023. "I wouldn't say offensive but...": Disability-Centered Perspectives on Large Language Models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 205–216.
- [44] Anujay Ghosh, Monalika Padma Reddy, Satwik Ram Kodandaram, Utku Uckun, Vikas Ashok, Xiaojun Bi, and IV Ramakrishnan. 2024. Screen Reading Enabled by Large Language Models. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–5.
- [45] Kate Glazko, Yusuf Mohammed, Ben Kosa, Venkatesh Potluri, and Jennifer Mankoff. 2024. Identifying and Improving Disability Bias in GPT-Based Resume Screening. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 687–700.
- [46] Kate S Glazko, Momona Yamagami, Aashaka Desai, Kelly Avery Mack, Venkatesh Potluri, Xuhai Xu, and Jennifer Mankoff. 2023. An Autoethnographic Case Study of Generative Artificial Intelligence's Utility for Accessibility. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 1–8.
- [47] Ricardo E Gonzalez Penuela, Jazmin Collins, Cynthia Bennett, and Shiri Azenkot. 2024. Investigating Use Cases of AI-Powered Scene Description Applications for Blind and Low Vision People. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–21.
- [48] Google. 2024. Gemini. Retrieved September 12, 2024 from <https://gemini.google.com/>
- [49] Google. 2024. Google Assistant. Retrieved January 10, 2024 from <https://assistant.google.com/>
- [50] Google. 2024. Google Workspace. Retrieved January 10, 2024 from <https://workspace.google.com/>
- [51] Accessible Technology Users Group. 2024. Accessible Technology Users Group. Retrieved March 22, 2024 from atug-subscribe@damorris.com
- [52] Yucheng Han, Chi Zhang, Xin Chen, Xu Yang, Zhibin Wang, Gang Yu, Bin Fu, and Hanwang Zhang. 2023. Chartllama: A multimodal llm for chart understanding and generation. *arXiv preprint arXiv:2311.16483* (2023).
- [53] Apple Inc. 2025. Accessibility - Vision. <https://www.apple.com/accessibility/vision/>
- [54] Nanna Inie, Jeanette Falk, and Steve Tanimoto. 2023. Designing participatory ai: Creative professionals' worries and expectations about generative ai. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–8.
- [55] Tatsuya Ishihara, Hironobu Takagi, Takashi Itoh, and Chieko Asakawa. 2006. Analyzing visual layout for a non-visual presentation-document interface. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. Association for Computing Machinery, New York, NY, USA, 165–172.
- [56] Youssef Bou Issa, Mustapha Mojahid, Bernard Oriola, and Nadine Vigouroux. 2009. Accessibility for the blind: An automated audio/tactile description of pictures in digital documents. In *2009 International Conference on Advances in Computational Tools for Engineering Applications*. IEEE, 591–594.
- [57] Chutian Jiang, Wentao Lei, Emily Kuang, Teng Han, and Mingming Fan. 2023. Understanding Strategies and Challenges of Conducting Daily Data Analysis (DDA) Among Blind and Low-vision People. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 1–15.
- [58] Martin Jung and Erdmuthe Meyer zu Bexten. 2002. Visually impaired persons using Microsoft Powerpoint presentations. In *Computers Helping People with Special Needs: 8th International Conference, ICCHP 2002 Linz, Austria, July 15–20, 2002 Proceedings 8*. Springer, 463–468.
- [59] Rie Kamikubo, Farnaz Zamiri Zeraati, Kyungjun Lee, and Hernisa Kacorri. 2024. AccessShare: Co-designing Data Access and Sharing with Blind People. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–16.
- [60] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacl-HLT*, Vol. 1. Minneapolis, Minnesota, 2.
- [61] Satwik Ram Kodandaram, Utku Uckun, Xiaojun Bi, IV Ramakrishnan, and Vikas Ashok. 2024. Enabling Uniform Computer Interaction Experience for Blind Users through Large Language Models. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) (ASSETS '24). Association for Computing Machinery, New York, NY, USA, Article 73, 14 pages. doi:10.1145/3663548.3675605
- [62] Gerd Kortemeyer. 2023. Using artificial-intelligence tools to make LaTeX content accessible to blind readers. *arXiv preprint arXiv:2306.02480* (2023).
- [63] Askat Kuzdeuov, Olzhas Mukayev, Shakhizat Nurgaliyev, Alisher Kumbolsyn, and Huseyin Atakan Varol. 2024. ChatGPT for visually impaired and blind. In *2024 International Conference on Artificial Intelligence in Information and Communication (ICAIC)*. IEEE, 722–727.
- [64] Cheuk Yin Phipson Lee, Zhuohao Zhang, Jaylin Herskovitz, JooYoung Seo, and Anhong Guo. 2022. Collabally: Accessible collaboration awareness in document editing. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–17.
- [65] Hae-Na Lee. 2023. *Enhancing the Usability of Computer Applications for People With Visual Impairments via UI Augmentation*. Ph. D. Dissertation. State University of New York at Stony Brook.
- [66] Hae-Na Lee, Vikas Ashok, and IV Ramakrishnan. 2020. Repurposing visual input modalities for blind users: a case study of word processors. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2714–2721.
- [67] Hae-Na Lee, Vikas Ashok, and IV Ramakrishnan. 2021. Bringing things closer: Enhancing low-vision interaction experience with office productivity applications. *Proceedings of the ACM on Human-computer Interaction* 5, EICS (2021), 1–18.
- [68] Barbara Leporini, Marina Buzzi, and Marion Hersh. 2021. Distance meetings during the covid-19 pandemic: are video conferencing tools accessible for blind people?. In *Proceedings of the 18th International Web for All Conference*. 1–10.
- [69] Barbara Leporini, Marina Buzzi, and Marion Hersh. 2023. Video Conferencing Tools: Comparative Study of the Experiences of Screen Reader Users and the Development of More Inclusive Design Guidelines. *ACM Transactions on Accessible Computing* 16, 1 (2023), 1–36.
- [70] Q Vera Liao and Jennifer Wortman Vaughan. 2023. Ai transparency in the age of llms: A human-centered research roadmap. *arXiv preprint arXiv:2306.01941* (2023), 5368–5393.
- [71] Irene Lopatovska, Katrina Rink, Ian Knight, Kieran Raines, Kevin Cosenza, Harriet Williams, Perachya Sorsche, David Hirsch, Qi Li, and Adrianna Martinez. 2019. Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science* 51, 4 (2019), 984–997.
- [72] Weizheng Lu, Jing Zhang, Ju Fan, Zihao Fu, Yueguo Chen, and Xiaoyong Du. 2024. Large language model for table processing: A survey. *arXiv preprint arXiv:2402.05121* (2024).
- [73] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 5286–5297.
- [74] Haley MacLeod, Cynthia L Bennett, Meredith Ringel Morris, and Edward Cutrell. 2017. Understanding blind people's experiences with computer-generated captions of social media images. In *proceedings of the 2017 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 5988–5999.
- [75] Kamil Malinka, Martin Peresini, Anton Firc, Ondrej Hujnák, and Filip Janus. 2023. On the educational impact of chatgpt: Is artificial intelligence ready to obtain a university degree?. In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*. 47–53.
- [76] Daniela Massiceti, Camilla Longden, Agnieszka Slowik, Samuel Wills, Martin Grayson, and Cecily Morrison. 2024. Explaining CLIP's performance disparities on data from blind/low vision users. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12172–12182.
- [77] Meta. 2025. Meta AI. <https://www.meta.ai/>
- [78] Microsoft. 2024. Microsoft 365. Retrieved January 10, 2024 from <https://www.microsoft.com/en-au/microsoft-365>
- [79] Microsoft. 2024. Microsoft 365 Copilot. Retrieved February 9, 2024 from <https://www.microsoft.com/en-us/microsoft-365/enterprise/microsoft-365-copilot>

- [80] Microsoft. 2024. Seeing AI. Retrieved September 12, 2024 from <https://www.microsoft.com/en-us/garage/wall-of-fame/seeing-ai/>
- [81] Microsoft. 2025. Accessibility—for everyone. <https://www.microsoft.com/en-au/windows/accessibility-features>
- [82] Lourdes Morales, Sonia M Arteaga, and Sri Kurniawan. 2013. Design guidelines of a tool to help blind authors independently format their word documents. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 31–36.
- [83] Lourdes M Morales, Sonia M Arteaga, Peter Cottrell, and Sri Kurniawan. 2014. Design guidelines of tools for facilitating blind people to independently format their documents. In *Computers Helping People with Special Needs: 14th International Conference, ICCHP 2014, Paris, France, July 9-11, 2014, Proceedings, Part I 14*. Springer, 634–641.
- [84] Giulio Mori, Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Victor MR Penichet. 2011. Making “Google Docs” user interface more accessible for blind people. In *Advances in New Technologies, Interactive Interfaces, and Communicability: First International Conference, ADNTIC 2010, Huerta Grande, Argentina, October 20-22, 2010, Revised Selected Papers 1*. Springer, 20–29.
- [85] Jakob Nielsen. 2024. AI: First New UI Paradigm in 60 Years. Retrieved September 6, 2024 from <https://www.nngroup.com/articles/ai-paradigm/>
- [86] National Federation of the Blind. 2024. BlindResearch – Blindness-related research discussion and collaboration list. Retrieved March 22, 2024 from https://nfbnet.org/mailman/listinfo/blindresearch_nfbnet.org
- [87] Ollama. 2025. Get up and running with large language models. <https://ollama.com/>
- [88] OpenAI. 2024. ChatGPT. Retrieved September 12, 2024 from <https://openai.com/chatgpt/>
- [89] Sassy Outwater-Wright. 2020. Putting an end to the idea of working twice as hard: How Leadership Can Affect Change in the Workplace for People Who Are Blind. *Journal of Visual Impairment & Blindness* 114, 1 (2020), 70–72.
- [90] Yi-Hao Peng, Peggy Chi, Anjuli Kamman, Meredith Ringel Morris, and Irfan Essa. 2023. Slide Gestalt: Automatic Structure Extraction in Slide Decks for Non-Visual Access. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14.
- [91] Yi-Hao Peng, Jason Wu, Jeffrey Bigham, and Amy Pavel. 2022. Diffscrifer: Describing Visual Design Changes to Support Mixed-Ability Collaborative Presentation Authoring. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–13.
- [92] Minoli Perera, Bongshin Lee, Eun Kyoung Choe, and Kim Marriott. 2024. Visual Cues for Data Analysis Features Amplify Challenges for Blind Spreadsheet Users. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–16.
- [93] Mahika Phutane, Crescentia Jung, Niu Chen, and Shiri Azenkot. 2023. Speaking with My Screen Reader: Using Audio Fictions to Explore Conversational Access to Interfaces. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 1–18.
- [94] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. “Accessibility Came by Accident” Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–13.
- [95] Aaron Preece. 2025. Enhancing digital Accessibility through AI. Retrieved February 5, 2025 from <https://www.afb.org/aw/fall2024/digital-accessibility-ai>
- [96] Samuel Reinders, Swamy Ananthanarayan, Matthew Butler, and Kim Marriott. 2023. Designing conversational multimodal 3D printed models with people who are blind. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2172–2188.
- [97] Samsung. 2025. Do more with Bixby. <https://www.samsung.com/us/apps/bixby/>
- [98] Sergio Sayago and Mireia Ribera. 2020. Apple Siri (input)+ Voice Over (output)= a de facto marriage: An exploratory case study with blind people. In *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion*. 6–10.
- [99] 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*. Association for Computing Machinery, New York, NY, USA, 1–19.
- [100] John Gerard Schoeberlein and Yuanqiong Wang. 2014. Usability Evaluation of an Accessible Collaborative Writing Prototype for Blind Users. *Journal of Usability Studies* 10, 1 (2014).
- [101] Freedom Scientific. 2023. JAWS® – Freedom Scientific. <https://www.freedomscientific.com/products/software/jaws/>
- [102] Freedom Scientific. 2024. Picture Smart AI. Retrieved September 6, 2024 from <https://blog.freedomscientific.com/picture-smart-ai-how-we-got-here/>
- [103] Md Shamsujjoha, Qinghua Lu, Dehai Zhao, and Liming Zhu. 2024. A taxonomy of multi-layered runtime guardrails for designing foundation model-based agents: Swiss cheese model for ai safety by design. *arXiv preprint arXiv:2408.02205* (2024).
- [104] Jorge Sasaki Resende Silva, Paula Christina Figueira Cardoso, Raphael Winckler de Bettio, Daniela Cardoso Tavares, Carlos Alberto Silva, William Massami Watanabe, and André Pimenta Freire. 2024. In-Page Navigation Aids for Screen-Reader Users with Automatic Topicalisation and Labelling. *ACM Transactions on Accessible Computing* (2024).
- [105] Jorge Sasaki Resende Silva, André Pimenta Freire, and Paula Christina Figueira Cardoso. 2022. When headers are not there: Design and user evaluation of an automatic topicalisation and labelling tool to aid the exploration of web documents by blind users. In *Proceedings of the 19th International Web for All Conference*. 1–11.
- [106] AM Silverman, LP Rosenblum, EC Bolander, CR Rhoads, and K Bleach. 2022. Technology and accommodations: Employment experiences of US adults who are blind, have low vision, or are deafblind. American Foundation for the Blind.
- [107] A. M. Silverman, S. J. Baguhn, M.-L. Vader, E. M. Romero, and C. H. P. So. 2025. *Empowering or Excluding: Expert Insights on Inclusive Artificial Intelligence for People With Disabilities*. White paper. American Foundation for the Blind. <http://www.afb.org/AIRsearch>
- [108] Marita Skjue, Asbjørn Følstad, and Petter Bae Brandtzaeg. 2023. The user experience of ChatGPT: findings from a questionnaire study of early users. In *Proceedings of the 5th international conference on conversational user interfaces*. 1–10.
- [109] Tony Stockman. 2004. The design and evaluation of auditory access to spreadsheets. In *ICAD 04-Tenth Meeting of the International Conference on Auditory Display*.
- [110] Tony Stockman, Christopher Frauenberger, and Greg Hind. 2005. Interactive sonification of spreadsheets. In *ICAD 05-Eleventh Meeting of the International Conference on Auditory Display*.
- [111] Vincent Stragier, Omar Seddati, and Thierry Dutoit. 2023. Developing an Interactive Agent for Blind and Visually Impaired People. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences*. 248–253.
- [112] Maartje ter Hoeve, Robert Sim, Elnaz Nouri, Adam Fournery, Maarten de Rijke, and Ryen W White. 2020. Conversations with documents: An exploration of document-centered assistance. In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*. 43–52.
- [113] Lida Theodorou, Daniela Massiceti, Luisa Zintgraf, Simone Stumpf, Cecily Morrison, Edward Cutrell, Matthew Tobias Harris, and Katja Hofmann. 2021. Disability-first dataset creation: lessons from constructing a dataset for teachable object recognition with blind and low vision data collectors. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–12.
- [114] Utku Uckun, Rohan Tumkur Suresh, Md Javedul Ferdous, Xiaojun Bi, IV Ramakrishnan, and Vikas Ashok. 2022. Taming User-Interface Heterogeneity with Uniform Overlays for Blind Users. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*. 212–222.
- [115] Alexandra Vtyurina, Adam Fournery, Meredith Ringel Morris, Leah Findlater, and Ryen W White. 2019. Verse: Bridging screen readers and voice assistants for enhanced eyes-free web search. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 414–426.
- [116] Ruotong Wang, Ruijia Cheng, Denae Ford, and Thomas Zimmermann. 2024. Investigating and designing for trust in ai-powered code generation tools. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 1475–1493.
- [117] Mirza Muhammad Waqar, Muhammad Aslam, and Muhammad Farhan. 2019. An intelligent and interactive interface to support symmetrical collaborative educational writing among visually impaired and sighted users. *Symmetry* 11, 2 (2019), 238.
- [118] Justin D Weisz, Jessica He, Michael Muller, Gabriela Hoefler, Rachel Miles, and Werner Geyer. 2024. Design Principles for Generative AI Applications. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–22.
- [119] Boming Xia, Qinghua Lu, Liming Zhu, and Zhenchang Xing. 2024. Towards ai safety: A taxonomy for ai system evaluation. *arXiv preprint arXiv:2404.05388* (2024).
- [120] Alyson Yin, Riya Sogani, Bruna Oewel, Karen Phan, Jaime Shin Young Park, Mia Ashley Yeo, Lindsay Ann Yazzolino, Karen Arcos, Ali Abdolrahmani, Emily Blank, et al. 2024. “Malicious” Pictorials: How Alt Text Matters to Screen Reader Users’ Experience of Image-Dense Media. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 1262–1274.
- [121] Zhuohao Zhang and Jacob O Wobbrock. 2023. A11yBoard: Making Digital Artboards Accessible to Blind and Low-Vision Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–17.
- [122] Zhuohao Jerry Zhang, Gene SH Kim, and Jacob O Wobbrock. 2023. Developing and Deploying a Real-World Solution for Accessible Slide Reading and Authoring for Blind Users. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, New York, NY, USA, 1–15.

A Breakdown of Issues Reported in the Survey

Table 5: No. of users reporting issues in productivity apps

Areas	Challenges	Email	Word processing	Video conferencing	Spreadsheets	Calendar	Cloud storage	Instant messaging	Presentation tools	Note taking tools	Whiteboard tools	Occurs in all apps	Total no. of respondents
Understanding Content	Understanding non-graphical content (textual content, tabular data, formulas and finding files)	1	9	3	12	0	2	0	8	0	0	2	26
	Understanding graphical data (images, graphics, charts and shared screens)	7	17	6	8	1	0	0	14	0	1	7	35
	Understanding formatting (font attributes or color changes)	1	8	0	4	0	0	0	2	0	0	1	10
	Other (Use proximity to convey information, comparing content in documents, accessing templates, understanding the speaker during video conferencing)	0	0	1	1	0	0	0	1	1	0	1	4
Accessing Controls	Finding and using controls (locating and navigating through controls and unlabelled controls)	2	7	10	6	2	3	3	7	0	0	7	31
	Keystrokes related issues (lack of standard or not working keystrokes, remembering too many keystrokes)	0	2	2	1	0	0	1	0	0	0	3	8
	Use of ribbon or other types of menus	0	4	1	4	0	1	0	3	0	0	3	7
	Other (inconsistent use of controls, use of inaccessible controls)	1	1	0	0	0	0	0	1	0	0	1	3
Authoring and Editing	Authoring (creating non-graphical and graphical content such as charts, and visually appealing content)	0	2	0	2	1	0	0	6	1	1	0	10
	Editing content	0	4	0	3	1	0	0	1	0	0	3	11
	Applying formatting	1	5	0	1	0	0	0	4	0	0	1	6
	Using built in support features (spell check, auto-correct)	0	4	0	1	0	0	0	0	0	0	1	6
	Requiring sighted assistance for reviewing, formatting, or editing content	0	3	0	1	0	0	0	1	1	0	0	3
Navigation	Navigating content (issues related to content structure and reading order)	1	8	1	3	0	1	0	1	0	0	1	13
	Navigating around app interface	0	3	2	2	0	1	2	3	0	0	0	5
	High level general comments about navigation	0	0	1	0	0	1	0	1	0	0	0	2
Collaboration	Understanding and tracking others changes and comments	0	9	0	2	0	0	0	0	0	0	0	9
	High level general comments about collaboration	0	3	2	4	0	0	0	1	0	1	0	8
SR-related Issues	SR not reading content	2	4	2	3	0	1	0	4	1	0	1	12
	SR losing focus	1	0	1	0	0	0	0	0	0	0	2	4
	Other (inconsistencies with types of SRs, lack of responsiveness, extra verbosity)	0	3	1	2	0	0	0	1	0	0	0	4