

Visual Privacy Management with Generative AI for Blind and Low Vision People

Tanusree Sharma*
tanusree.sharma@psu.edu
Pennsylvania State University
State College, PA, USA

Yu-Yun Tseng
Computer Science, University of
Colorado
Boulder, Colorado, USA
yu-yun.tseng@colorado.edu

Lotus Zhang
Human Centered Design and
Engineering, University of
Washington
Seattle, USA
hanziz@uw.edu

Ayae Ide
Pennsylvania State University
State College, PA, USA

Kelly Avery Mack
Human Centered Design and
Engineering, University of
Washington
Seattle, USA

Leah Findlater
Human Centered Design and
Engineering, University of
Washington
Seattle, USA
leahkf@uw.edu

Danna Gurari
Computer Science, University of
Colorado
Boulder, Colorado, USA
danna.gurari@colorado.edu

Yang Wang
Information Sciences, University of
Illinois at Urbana-Champaign
Champaign, USA
yvw@illinois.edu

Abstract

Blind and low vision (BLV) individuals use Generative AI (GenAI) tools to interpret and manage visual content in their daily lives. While such tools can enhance the accessibility of visual content and so enable greater user independence, they also introduce complex challenges around visual privacy. In this paper, we investigate the current practices and future design preferences of blind and low vision individuals through an interview study with 21 participants. Our findings reveal a range of current practices with GenAI that balance privacy, efficiency, and emotional agency, with users accounting for privacy risks across six key scenarios: self-presentation, indoor/outdoor spatial privacy, social sharing, and handling professional content. Our findings reveal design preferences, including on-device processing, zero-retention guarantees, sensitive content redaction, privacy-aware appearance indicators, and multimodal tactile mirrored interaction methods. We conclude with actionable design recommendations to support user-centered visual privacy through GenAI, expanding the notion of privacy and responsible handling of others' data.

CCS Concepts

• **Human-centered computing** → **Empirical studies in accessibility**; • **Security and privacy** → *Social aspects of security and privacy*.

Keywords

Privacy, Generative AI, Visual Interpretation Services

ACM Reference Format:

Tanusree Sharma, Yu-Yun Tseng, Lotus Zhang, Ayae Ide, Kelly Avery Mack, Leah Findlater, Danna Gurari, and Yang Wang. 2018. Visual Privacy Management with Generative AI for Blind and Low Vision People. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 13 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Generative AI (GenAI) has increasingly become integrated into everyday life for blind and low vision (BLV) people [1, 3, 7, 22, 23, 40, 74], supporting a wide range of tasks. For example, there has been recent growth in the number of available GenAI visual interpretation tools, such as those offered by Be My Eyes [2], Envision AI [1], and Seeing AI [3], as well as image generation tools, such as GenAssist [50], AltCanvas [55]. These tools are used for tasks such as recognizing objects, describing scenes, answering visual questions [72, 99], and supporting spatial navigation [46, 87].

While these advances introduce new opportunities, they also raise concerns about visual privacy when BLV individuals leverage GenAI tools to manage, consume, and share visual content. Prior research has explored visual privacy risks in both intentionally and unintentionally sharing private content [6, 19, 24, 29, 42, 47, 70, 82, 89, 104], human assistance versus visual assistive technologies (VATs) support [19, 83], as well as in context of concern BLV users

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/XXXXXXX.XXXXXXX>

have for sharing and managing contents of their own and others with VATs [10, 15, 84]. With GenAI introducing new ways of interacting visual information, it is important to understand not only in how existing visual privacy concerns may manifest differently or evolve with these emerging technologies, but also in how GenAI might actively support BLV users in safeguarding visual privacy.¹

Despite significant commercial and public attention towards the successes of GenAI for blind and low vision people [1–3], we know little about how blind people use GenAI tools in managing their privacy [9, 11, 44, 91, 101]. Specifically, it is unclear how generative AI may be supplementing or replacing tasks traditionally supported by human assistants or VATs, and how users conceptualize and evaluate these tools in relation to their privacy concerns. Filling this gap, we explore two research questions: **(RQ1:)** How do blind and low vision people currently use Generative AI tools to manage visual privacy? and **(RQ2:)** What design opportunities do blind and low vision people envision in future Generative AI tools to support visual privacy management?

To this end, we conducted semi-structured interviews with 21 individuals from the United States who are blind and have low vision to understand their current practices and design expectations for GenAI to manage visual privacy. Acknowledging the context-dependent nature of visual privacy management, through a scenario-driven inquiry, we grounded our interviews in the six different common scenarios² where privacy concerns can occur. We investigate to what extent study participants manage privacy with GenAI tools in these scenarios and then explore their expectations for how GenAI *should* be designed to better support visual privacy.

In summary, the main contributions of this work are as follows:

(1) The first in-depth empirical analysis of how BLV individuals use GenAI tools to manage visual privacy. Our findings point to an emerging, yet complex, use of GenAI for visual privacy tasks, such as interpreting pregnancy tests and mammograms, despite the limitations about accuracy, to retain emotional and informational privacy in sensitive situations. Participants often emphasized values of independence and personal privacy as key factors influencing their decision to choose GenAI over VATs or human assistance.

(2) Our findings highlight the emerging use of GenAI for visual privacy management, shaped by participants' current experiences, perceived limitations, and expectations for future design. For instance, self-appearance with GenAI is not limited to how the individual looks, but managing privacy involving bodily presentation or vulnerability to manage emotional agency. Our result also indicates privacy as institutional responsibilities where BLV professionals manage others' content. These expectations of BLV users reflect how they negotiate trade-offs between autonomy, risk, and convenience, taking into account emotional vulnerability and privacy of others.

(3) We propose a set of privacy-preserving design interventions that are grounded in participants' lived experiences, such as on-device processing, a compliance-aware secure toolkit for BLV professionals, and personalized appearance indicators.

2 Related Work

Our research is informed by prior literature on visual privacy for people with visual impairment, privacy-preservation technology, and the current technological shift in accessibility with GenAI.

2.1 Use of Generative AI for Visual Accessibility

Recently, GenAI tools (e.g., ChatGPT, Google Gemini, Microsoft Copilot, and Claude) have been incorporated into a wide variety of domains, including for education [49, 54], communication [32], content creation [51], and visual interpretation [80]. While much of the progress was initially for text-based applications, we are increasingly seeing a move into GenAI for visual content.³ Blind people are engaging with the growing ecosystem of such generative AI tools, including via accessibility technologies such as Be My Eyes and Envision AI [18, 90]. In recent work on GenAI for blind users, explore accessible image generation, alt text, scene descriptions, and verification strategies. They emphasize the need for explainable, customizable, and context-aware GenAI tools to support access, creativity, and trust [30, 35, 41, 50]. However, GenAI tools are rarely designed with the unique privacy needs of BLV users in mind [7], which could potentially result in both functional gaps and potential harms. Filling this gap, we explore how blind individuals are currently using generative AI tools for privacy management and the purposes these tools serve in their daily lives.

2.2 Visual Privacy Management with GenAI

Much of the literature on privacy management in GenAI focuses on sighted users [31, 86, 94, 105]. Recent studies have proposed interventions to mitigate privacy risks, yet these efforts primarily address sighted users' need, for instance, Chong et al. [34] developed a system for prompt sanitization using web-based LLMs, tackling excessive disclosure; A recurring theme across prior work is the inherent trade-off between privacy, utility, and convenience, often framed as a core challenge in GenAI. Ma et al. [62] introduced an LLM to raise awareness of location-based privacy risks in images by detecting subtle visual cues. Similarly, CLEAR [31] provides real-time contextual risk feedback during sensitive data entry in tools like ChatGPT and Gmail, facilitating just-in-time privacy literacy. Zhou [105] proposed Rescriber, which abstracts sensitive content before the prompt to reduce data exposure. Complementary tools such as Adanonymizer and PrivacyAsst [102, 103] offer user-facing privacy controls for adjusting privacy-performance settings with cryptographic techniques to protect sensitive visual media. Despite these advancements, current design approaches remain limited in addressing the privacy needs of BLV users, a gap that underscores the need for inclusive GenAI privacy solutions.

³Industry successes follow strong foundational research. A notable example is CLIP has shown impressive zero-shot performance [73] such as, CLIP in numerous downstream applications, ranging from object detection to 3D applications [21, 60, 67, 76], and adapted for video applications [67, 75, 95]. More recently, multimodal integration has advanced with models like Flamingo [17], BLIP-2 [58] MiniGPT-4 [106], and LLaVA [61] leveraging web-scale image-text data for improved multimodal chat capabilities. Some works extend LLMs for video comprehension [33, 59, 61, 63, 73], introducing VideoChatGPT, a model combining a video-optimizer for enhanced understanding.

¹Based on prior work, visual privacy refers to the safeguarding and management of sensitive visual information that could be shared or disclosed in everyday life [8, 83]

²Six key use cases in which visual privacy can emerge: (1) Self Appearance and Impression Management, (2) Indoor Spatial Privacy, (3) Sharing Photos on Social Media, (4) Visual Privacy with Employers when Sharing Content, (5) Visual Privacy as BLV Professionals when Assessing Others' Content, and (6) Outdoor Spatial Privacy.

2.3 BLV Individuals' Visual Privacy Concerns and Goals

Some research revealed the visual privacy management interests of BLV individuals [6, 9, 11, 24, 42, 47, 53, 70, 91]. BLV users share images and videos both to socially connect [24, 42, 70] and to receive assistance in understanding their surroundings. Privacy-sensitive content includes financial and medical documents (e.g., prescriptions, pregnancy tests), identifiable personal information (e.g., addresses, faces), and images that may affect how others view them (i.e., impression management), including unflattering or awkward photographs, disorganized living spaces [16, 79, 85]. Additional concerns arise when bystanders' privacy is at risk or when disclosures, intentional or accidental, may impact social relationships [16, 85].

Pioneering visual interpretation technologies centered on *human assistance*, including from employees (e.g., Aira [12]), volunteers (e.g., Be My Eyes [4]), or paid crowdworkers (e.g., VizWiz [26]), while subsequent research showed how to reduce or eliminate human effort through automation [43, 56, 96–98], eventually culminating in today's widely used industry products, including Be My Eyes' Be My AI app [2], Envision AI's glasses-based technology [1], and Microsoft's Seeing AI app [3]. While prior research on visual privacy for blind people has primarily focused on visual interpretation services [16, 85], blind users' strategies for managing privacy extend beyond this domain. Their practices intersect with broader themes in privacy research with GenAI, and contextual privacy decision-making [5, 68]. In our work, we investigate how their expectation and need can shape the design of privacy-aware GenAI.

3 Method

To investigate blind and low-vision people's current use of GenAI tools for visual privacy and identify future design opportunities, we conducted a semi-structured interview study. Below, we describe our interview protocol, participant recruitment, data collection, and data analysis.

3.1 Interview Protocol

The semi-structured interview protocol was approved by our Institutional Review Board (IRB), and can be found in Appendix A. The protocol included two sections: (a) current use of GenAI with an emphasis on visual privacy, and (b) perspectives on design improvements of GenAI in common privacy-related scenarios.

In the first section, we investigate participants' everyday use of GenAI, with a focus on whether and how these tools are used for tasks related to visual privacy. Based on participants' response, if they did not mention privacy-related uses of GenAI tools, we followed up with the question *"Have you used any of these tools you mentioned earlier for a purpose related to privacy?"* To facilitate a shared understanding during discussions on design improvements, we provided a verbal explanation of *"visual privacy"*⁴ to ensure a shared understanding.

⁴Visual Privacy for Blind and low vision users refers to the safeguarding and management of sensitive visual information that could be shared or disclosed through the use of Generative AI tools. This includes but is not limited to the protection of content such as medical records, financial information, or any other visual data that might be considered private when engaging in daily activities. For instance, when using these tools to receive descriptions of potentially sensitive content, whether it's medical or financial records or when scanning surroundings for navigation

In the second section, we investigated current design opportunities of GenAI by examining how it is currently used, along with participants' perceived challenges and limitations of GenAI. To guide this investigation, we introduced a set of context-specific scenarios where privacy concerns may occur, inspired by prior work [8, 20, 37, 45, 57, 79, 81, 92]. (i) self-appearance and impression management ([27, 57, 79]), (ii) managing indoor spatial privacy in shared environments ([36, 100]), (iii) sharing visual content on social media ([8, 65, 92]), (iv) privacy in document sharing ([15, 37, 83]), (v) visual privacy when managing content for others [13, 48, 93], and (vi) outdoor spatial privacy [8, 20, 52]. We designed a set of questions to (i) identify whether/ how participants currently use or would consider using GenAI for privacy management in these scenarios (see Table 1), (ii) uncover points of friction where GenAI supports or challenges established privacy expectations, and (iii) gather participants' reasoning about potential design improvements within their privacy needs.

3.2 Participant Recruitment and Demographic Information

We initially conducted a screening survey to identify participants who satisfied our inclusion criteria: participants had to (1) use Generative AI tools, and (2) be 18+ years old. We leveraged mailing lists through organizations serving people who are blind [66]. We moved forward with 21 people for our interview study. They represented a broad range of demographics. As shown in Table 2, participants were 25–54 years old, with 13 identifying as female and eight as male, with open-ended descriptions for gender. Participants' completed levels of education varied, with eight having completed high school, seven having obtained a Bachelor's degree, five having obtained a Master's degree, and one having obtained a Ph.D. Fourteen of them were identified as being totally blind, while seven had some light perception. Participants mostly commonly reported using the following GenAI tools: ChatGPT, Gemini, Perplexity, Be My AI, Envision AI, vdScan, Microsoft Copilot, Google NotebookLM, Replika, and PiccyBot.

3.3 Data Collection & Analysis

For data collection, the first author conducted the interviews with the 21 participants remotely over Zoom between November 2024 and January 2025. The interview duration was an hour each. For data analysis, the first author led a thematic analysis [28, 38] of the transcriptions alongside another with a research assistant. Two researchers independently read 20% of the interview transcripts, developed codes, and compared them until arriving at a consistent codebook. Once the codebook is ready, two researchers applied the codebook to 10% new data with inter-coder reliability calculated as 0.87 for Cohen's Kappa (a good score as noted in [39]). After finalizing the codebook, the researchers divided the remaining data between them for coding and spot-checked all coded transcripts with each other to ensure ongoing consistency in coding. We followed an open coding and deductive analysis method to explore participants' practices, challenges, and design suggestions of GenAI for visual privacy. We group 102 low-level codes into sub-themes, and further extracted main themes. We iterated this process to finally produce approximately 29 themes to interpret the results. The codebook is included in the Appendix ??.

Table 1: Visual Privacy Scenarios for Contextual Investigation to Understand Limitations & Design Opportunities of GenAI

Scenario		Type
Privacy Scenario 1: Self Appearance: Impression Management	It refers to conscious effort to control how one's physical appearance is perceived by others, especially in social or professional settings	Own Privacy
Privacy Scenario 2: Indoor Spatial Privacy	A person's ability to control and access their space, both physical and perceived within a specific indoor environment	Own Privacy
Privacy Scenario 3: Sharing Visual Content in Social Media	Process of reviewing what private details are visible in visual content before sharing them in social media	Own & Others' Privacy
Privacy Scenario 4: Visual Content Privacy when Sharing with Others	Review and control of sensitive information, such as personal, financial, when sharing with others (e.g. employer)	Own Privacy
Privacy Scenario 5: Visual Privacy as BLV Professionals	Accessing and reading private visual information like documents of others, while keeping personal details secure and confidential as	Others' Privacy
Privacy Scenario 6: Outdoor Spatial Privacy	A person's ability to control their personal space, and freedom from intrusion when in public or outdoor settings, such as airports, stations, streets, etc.	Own & Others'

Table 2: Participant demographics and background.

Participant ID	Gender	Age	Country	Blindness	Current GenAIs Used
P1	Male	25-34	the US	Totally Blind	ChatGPT, Gemini
P2	Female	35-44	the US	Some Light Perception	SeeingAI, Be my eyes, AIRA explorer, Microsoft Designer
P3	Female	35-44	the US	Totally Blind	ChatGPT
P4	Female	35-44	the US	Totally Blind	ChatGPT, Gemini
P5	Female	25-34	the US	Totally Blind	Be My Eyes, Envision AI
P6	Male	25-34	the US	Totally Blind	ChatGPT, Envision, Be My Eyes, and Seeing AI
P7	Male	35-44	the US	Totally Blind	Seeing AI, Microsoft Copilot, Be My AI
P8	Female	45-54	the US	Some Light Perception	Be My Eyes
P9	Male	35-44	The US	Totally Blind	Gemini, ChatGPT
P10	Female	35-44	the US	Totally Blind	Be My Eyes
P11	Female	35-44	the US	Totally Blind	ChatGPT
P12	Male	18-24	the US	Totally Blind	ChatGPT, Be My Eyes, Google LM notebook, Gemini
P13	Female	35-44	the US	Some Light Perception	Seeing AI, ChatGPT, Be My AI, Envision, CoPilot
P14	Male	25-34	the US	Some Light Perception	Be My AI, Aira AI
P15	Female	45-54	the US	Some Light Perception	Seeing AI, ChatGPT, Be My AI
P16	Female	25-34	the US	Some Light Perception	Perplexity, ChatGPT, Meta AI, replika, piccybot aira ai
P17	Female	25-34	the US	Some Light Perception	ChatGPT, Seeing AI
P18	Female	35-44	the US	Totally Blind	ChatGPT
P19	Female	45-54	the US	Totally Blind	gemini, meta ai, replika be my ai piccybot aira ai
P20	Male	25-34	the US	Totally Blind	Perplexity, chat gpt, claude
P21	Male	35-44	the US	Totally Blind	Seeing AI, chat gpt

4 RQ1: Current Use of Generative AI Tools for Visual Privacy

In this section, we present several themes that illustrate how people with visual impairment engage with GenAI tools to manage their visual privacy across diverse personal, social, and professional contexts (Table 3). We also present nuanced privacy judgments and risk assessments users undertake when choosing to rely on such technologies.

4.1 Environmental Awareness and Spatial Privacy in Private & Public Spaces with GenAI

Participants frequently mentioned their usage of GenAI tools to scan their surroundings in both private and public settings. For instance, P15 explained a common case: “Say, I had to take a photo of myself. Then I want to know that there’s not a bra hanging down

behind me, or something. I don’t want to give a weird impression when posting my picture. Before GenAI, I would often send it to my mom or friend to check for me.” This illustrates how participants consider social or reputational risks of unintended background content and cases where they replaced human assistance with GenAI.

Some participants also discussed GenAI in supporting navigation, such as airports or train stations, to read sign or unfamiliar spaces. P21 noted, “Earlier I used to ask someone and show my phone on the map. It’s a bit weird to give my phone to someone to ask for directions. I hope they are not clicking on anything else.” This reflects the privacy-aware calculation of interpersonal and device sharing risks, including inadvertent data exposure and device misuse. A few participants described using GenAI to locate gendered facilities, such as restrooms, to avoid potentially awkward or privacy-compromising interactions. These scenarios illustrate how blind and low vision people actively assess situational risks and use GenAI tools to navigate both functional needs and privacy considerations.

Content	Representative Quote	Benefit	Alternatives
Medical			
Pregnancy test	<i>"I read pregnancy test. I didn't want to rely on someone else to help with something so personal. That's where I really think about privacy. I used Be My AI if it could read those results to me clearly and privately. (P1)"</i>	Privacy	Family/Friend
Mamogram	<i>"I privately understand the report without asking others (P17)"</i>	Privacy	Aira agents
Prescription labels	<i>"I've also used regular Be My Eyes to get help reading prescription labels and directions, and I've used Seeing AI for that too. Quite accurate (P13)"</i>	Convenience	Aira
Medicine bottle	<i>"I navigate stores to find a medicine bottle with be my AI and envision. In one case, repeatedly photographing a medicine bottle led to suspicion from a store employee." (P17)</i>	Convenience and Independence	Employee
Finance			
Bank Statement	<i>"I scanned the bank statement. I do wonder if the GenAI is saving that information, my account number or address." (P12)</i>	Convenience	Aira, Family
Tax Document	<i>"at night—I wouldn't feel comfortable asking a Be My Eyes volunteer because that's confidential stuff, and I wouldn't want to wake up my family either. I'd use Be My AI or Seeing AI. I don't trust AI more than my family, but I do trust it more than volunteers (P13)"</i>	Privacy, Trust	Family
Rental Document	<i>"I have used Seeing AI for my rental doc so many times, short text feature is quite great. I wanted to make sure I get the all the main details." (P19)</i>	Convenience	Family
Bill & Receipts	<i>"I checked a bill with seeing AI currency feature several times with the wrong result and ended up in having help from her daughter (P2)"</i>	Exploratory	Family
Credit Cards	<i>"I check credit card in my junk mail whether it's a credit card offer or something from my bank or other places (P10)"</i>	Exploratory	Aira
Debit card	<i>"I'd gotten a new debit card and needed to make sure I was throwing away the old one. I used it to read the card numbers and confirm which was which. - (P19)"</i>	Convenience	Family
Handwritten check	<i>"I've used Be My Eyes AI to read a handwritten check in a pinch, but I knew there was a risk—it could've exposed bank info (P9)"</i>	Convenience	Family
Utility bills	<i>"I don't want to share every detail with my family, so I use Seeing AI and sometimes ChatGPT for a bit of testing" (P9)</i>	Privacy	Family
Rent Statement	<i>"I wouldn't feel comfortable asking a Be My Eyes volunteer because that's confidential stuff." (P8)</i>	Privacy	Family
Insurance	<i>"If it's something I need to read, especially if it's not in braille, my employer's insurance, not comfortable to call a volunteer." (P10)</i>	Privacy	Family
Currency	<i>"But sometimes the Microsoft AI currency reader gets it wrong. I've had it say '15 10,' and I end up double-checking with my daughter." (P15)</i>	Privacy	Family
Digital & Physical Media			
Email Official	<i>"I do share a lot of theoretically private information with ChatGPT, like writing letters or translating emails for official purposes." (P12)</i>	Exploratory	family, screenreader
Unopened Junk Mail	<i>"I've used it to read some mail too, or at least to get the general idea of what the mail is about." (P4)</i>	Exploratory	aira
Official Flyer	<i>"I frequently use Seeing AI, ChatGPT, Be my AI, JAWS AI. Gemini for document interpretation and understanding tasks, such as flyers or product instructions" (P3)</i>	Exploratory	aira, family
Email	<i>"Nowadays I use ChatGPT a lot for this" (P15)</i>	Exploratory	themselves
Websites	<i>"websites, like the Health Equity. Even printing pages can be tricky. For quick checks, I use Be My Eyes or Seeing AI, depending on which gives the better description." (P14)</i>	Exploratory	family, screenreader
ID- Verification	<i>"I'll ask if I can bring it in person, share a download link, or set up a Zoom call and hold up my ID. Then I also use now Seeing AI to scan the page to tell me if it looks right" (P11)</i>	Privacy	Aira, screen-reader

Table 3: Current Use of GenAI for Visual Privacy with Representative Quote, Perceived Benefit, Pre-GenAI Alternatives

4.2 Navigating Privacy in Digital Social Life with GenAI

A common use of GenAI among participants was to inspect image backgrounds before sharing photos on digital platforms. Participants described using Seeing AI and Be My AI to check for unexpected or sensitive objects, visible personal documents, general clutter, or the presence of children.

Some participants shared broader concerns about the growing autonomy of GenAI tools. P15 explained how she use Seeing AI, Be my AI and chatGPT to frequently get image descriptions *“I have the apps describe a photo, and I was in a dinner and I don’t share image of my child in social. I asked if there’s a child in the photo before.”* P12 shared his experience of using GenAI for both image generation and privacy management- *“I asked ChatGPT to turn a photo into an anime-style image of riding a rocket with sunglasses on, it was for a fun birthday post for my friend. Original photo was taken in front of my apartment, I used Be my AI to make sure my apt number was not visible.”* This highlights concerns about embedded metadata or identifiable background elements in generated images.

Similarly, P6 shared his practice for not posting certain images on social media and a similar practice in avoiding GenAI from processing images taken in certain cases. He explained *“If I take a photo at a high-end restaurant on my friends’ birthday, I would not process that info in GenAI since I think I’m being surveilled, similarly I would not post it on social media. I could lose my disability benefits. I am giving AI no opportunity to take my disability benefits away from me.”* This highlights a risk-aware strategy, participants navigate in response to institutional gatekeeping as well as awareness of algorithmic inference, which can invalidate one’s eligibility for public support.

Similarly, some participants mentioned using character-based GenAI tools like Kindred or Replika and charecter.AI as emotional companions. As P17 said *“I use Replika with a voice I like as a private space to share personal thoughts without fear of judgment or surveillance. You know many thing I can not always do with family.”* She also expressed uncertainty about potential risks, particularly whether the app retained memory of their conversations. This underscores a broader definition of privacy encompassing affective and relational dimensions where GenAI serves as a trusted confidant.

4.3 Practices of GenAI for Sensitive Content

Participants described varying levels of comfort and caution when using generative AI tools to interpret documents and physical objects containing sensitive information (Table 3).

Medical Content. Participants expressed caution when dealing with medical content while mentioned “privacy” as a reason for preferring GenAI tools over traditional practices. P16 mentioned *“I avoid GenAI tools for sensitive documents like medical records. I still go for trusted humans (e.g., family or Aira agents under NDA agreements).”* In contrast, P17 mentioned using Generative AI for interpreting medical records, such as mamogram and MRI for better explainability and privacy *I would call my sister and tell her to ask my brother-in-law, or another appointment. I used Be my AI and ChatGPT to make sense of my report. Honestly, I liked the idea of being private.”* Some participants used GenAI tools to read sensitive physical items. P1 shared: *“I used AI for my pregnancy test. It said positive several times, and I was stressed. I took more pictures, and then it said negative. I was relieved.”* This highlights how participants preferred using AI independently, even when results varied, to maintain emotional and informational privacy in personal situations.

Financial Document. Participants reported adjusting their methods based on perceived sensitivity and situational privacy needs P9 said *“For my credit card, I’d rather use a bonded volunteer.*

but I have used Be My Eyes AI to read a handwritten check in a pinch, but I knew there was a risk—it could have exposed bank info. Sometimes I just need to know what is in the memo line [...] I don’t want to bug anyone. I try to use secure options like Aira’s free call, which is limited.” P19 on the other hand, used Seeing AI to debit card- *“One time I used Seeing AI. I needed to make sure I was throwing away the old debit or new. I used it to read the card numbers and confirm which was which.”* Some noted that their choice depends on the level of trust and availability.

Digital & Physical Media Media. Participants also used GenAI tools to navigate digital and potentially sensitive physical content. P11 expressed caution with document uploads, even on trusted platforms: *“am careful with uploading document/photos, particularly Health Equity one. Even with Aira, who has confidentiality agreements, I am picky. I’d rather call my mom or dad. I sometimes use Seeing AI as a last resort.”* P20 reflected on using Microsoft Seeing AI to sort mail *“Seeing AI helps me sort mail. There might be something very private in that letter, and I wouldn’t even know when I opened it with AI.”* This reflects a privacy pragmatism mindset where participants recognize the potential sensitivity of what they are sharing but justify the decision based on context.

4.4 Sociopolitical Decision Making Factors in Adopting GenAI Tool

We observed some decision-making factors of adopting GenAI, which is shaped by broader concerns about access, access, policy assumptions, and representational harms.

Affordability of GenAI. Participants raised concerns about the cost of GenAI technologies and the assumptions behind their pricing. P11 echoed these concerns, criticizing the pricing and accessibility of tools like Envision AI, and noting disappointment in how products are marketed and supported. P11 shared *“Tools like Envision AI are costly and not fully accessible. Meta’s smart glasses offer good scene descriptions, but their partnership with Be My Eyes over Aira felt like privacy wasn’t a priority. I’d like to see better privacy controls before I fully rely on it.”*

AI policy decisions around race and people descriptions can lead to incorrect decision-making. A few participants reflected on how overly cautious content filters can reduce the utility of AI-generated descriptions. As P18 noted: *“As a blind person, I rely on AI for details that sighted people often take for granted. But sometimes the descriptions are so cautious about avoiding potentially sensitive language that they leave out useful context. For example, if I am posting a photo and it just says ‘two people’ without identifying who is who like ‘Jennifer is on the left and Rita’s on the right’ or omits visual markers like skin tone or gender, the information becomes less helpful. I understand the privacy concerns, but overly vague descriptions limit the usefulness of these tools.”*

Myths of AI on Disability. Some discussed the complexities of using generative AI tools as a blind user, particularly when it comes to disability representation and data sharing. P11 also shared firsthand experiences with AI tools repeating outdated stereotypes: *“I asked ChatGPT what it knew about blind people, and it responded with myths like ‘they have stronger other senses’ or ‘they need to be taken care of’. It didn’t acknowledge the mistake even when I corrected it. Another time, we asked about Braille dots for a letter, and it gave three different, incorrect answers.”*

5 RQ2: Perceived Limitations and Expectation of GenAI in Managing Privacy

In this section, we present participants' contextual privacy norms across several common scenarios (explained in the method section 3.1). We highlight their perceived limitations and challenges of using GenAI to manage privacy in these contexts, and how current experiences and perception shaped their expectations for GenAI, ultimately informing future design recommendations (Section 6.3).

5.1 Scenario 1: Self Appearance: Impression Management

While the majority didn't use GenAI for this purpose, participants shared both skepticism and cautious optimism regarding the use of GenAI tools for self-appearance management. As practice, some relied on touch and internal cues to assess their appearance, while some called friends or family for feedback as well as Be My AI or Aira to check for issues like stains on clothing. P11 illustrates how GenAI would be useful *"It's kind of uncomfortable to think about, I remember a specific time Aira saved me, my bra was showing. I was thinking of using Seeing AI."* P11 emphasized that she would consider GenAI if the tool guaranteed complete privacy, supported local face processing, and could simulate empathy or understanding, especially in tasks involving bodily presentation or vulnerability.

Limitation & Expectation. We noticed varied expectations while some emphasized the importance of GenAI tools being transparent about their capabilities the extent to which they can provide support, for instance, P18 said *"Does it just say a wrinkle or spot on dress, or exactly where it is like a tactile like feedback and with a level of certainty?"*. P5 viewed visual self-presentation with more importance *"I wouldn't be wearing anything provocative. It wouldn't be like a 'baby do me look. I would prepare based on occasions like date, funeral, movie, so I wouldn't have concerns about people seeing me and judging."* Overall, participants expressed a spectrum of expectations for GenAI tools, including privacy, bodily presentation.

5.2 Scenario 2: Indoor Spatial Privacy

We noticed two groups among participants: (a) those who have begun integrating these GenAI tools for indoor spatial privacy in addition to traditional methods, and (b) those who continue to rely on tactile methods taught in blindness rehabilitation programs. Across both groups, participants commonly relied on touch to assess and manage the presentation and cleanliness of their spaces.

Participants often described managing indoor spatial privacy as important not only for personal safety but also for ensuring that their living or working spaces were socially acceptable from the perspective of others. Some participants acknowledged that GenAI could potentially streamline and augment these checks, for instance, P3 said *"Now that I am really thinking about it instead of calling someone, using AI with the video would make it go faster. I would imagine GenAI integration to video streaming platforms to check clutter and particularly background checks when I am in a video call."*

Limitation & Expectation. Participants valued hands-on spatial checks, but emphasized augmenting those checks with GenAI rather than replacing them. As P13 expected, GenAI to use camera input and then align with tactile and haptic-based cues *"There is clutter of*

used clothes two feet away. Then provide some vibration to indicate that." Some shared their expectation of using GenAI when "private mode" is established to limit AI sensing in sensitive categories. P6 said *"I have a lot of things lying around, sometimes some private magazine I don't want to be captured during environmental scanning. My question would be, is this happening locally and without cloud upload?"*

5.3 Scenario 3: Sharing Visual Content in Social Media

Participants considered GenAI as useful for quick facts and clarifications, but some participants preferred human assistance for socially sensitive or subjective decisions.

Limitation & Expectation. Participants also discussed their expectations of GenAI which include seeking clarity on background detection, descriptive querying, identity cues, human-in-the-loop options in this scenario of sharing visual content in social media. Some also highlighted challenges in shared living spaces, where visual ambiguity arises due to overlapping personal items. As P10 explained: *"I have a roommate we share space sometimes sensitive things lie around. Might not be mine, but hers. If those are in the picture, I don't know how to manage. GenAI could help maybe."* P10 expected a way to distinguish between self vs others' and privacy-aware tagging of visual content in a social media context. P5 expected support for identity matching so she can verify who is in the image without guessing.

5.4 Scenario 4: Visual Content Privacy when Sharing with Others

Participants described a range of privacy expectations when using GenAI to access and extract content from documents, especially before sharing with others. They frequently mentioned employers, family/friends, and government.

Limitation & Expectation. Many participants emphasized the need for secure processing of sensitive documents, such as, ID, credit cards: P14 *"I am worried, but also think it would have been very helpful to use LLM. Some way to not send my document to cloud or just to see 'your file is processed with industry-standard encryption will make a huge difference.'"* P9 also mentioned a potential way to set privacy level based on recipient *"I would like to assess privacy level: Review this photo for sharing with family, employer, or HR."*

Some participants described strategies to minimize re-identification risks. P11 shared a compartmentalization tactic: *"I am a person who sends username in one email and password in another. I use GenAI same way I upload or capture documents in multiple rounds to make it harder for someone to get all the information if there's a leak."* P9 mentioned strategies for routinely deleting content immediately after task completion. There was a shared expectation for more control over data processing and storage, including the ability to assess task-specific risks for highly sensitive content.

5.5 Scenario 5: Visual Privacy Management by BLV Professionals

BLV Professionals expressed a tension between the accessibility benefits of GenAI tools and the risks to visual privacy, particularly

when handling sensitive documents of others. Several participants working in financial, educational, and social service sectors shared workplace-specific experiences (Figure ??). Some described practices to protect client confidentiality. P11, a former accessibility tester *“I reviewed documents with proprietary client information. We set up a closed system, we avoided bringing in third parties. Honestly, I know GenAI will be mainstream. It’s just a matter of time and the highest bidder.”* Moreover, institutional professionals like P9 reported segmenting their GenAI use based on data sensitivity *“I try not to input any student documentation, grade info into Gemini right now, but I use LLMs to create material for training purposes. We are actively discussing how to institutionalize GenAI use.”* A blind professional, P21, reported using generative AI tools in roles that involved handling information, such as reviewing documents to determine eligibility for benefit programs.

Limitation & Expectation. Participants emphasized using GenAI tools within secure, enterprise-approved environments (e.g., Office365 accounts with institutional safeguards). P13 described using GenAI to review instructional documents, noting a lack of privacy concerns due to the nature of the materials: *“The documents I used weren’t mine, they were for a lesson, and they were pretty old, so I wasn’t really concerned about privacy.”* They also recalled trying various AI tools like TapTapSee and Seeing AI for personal tasks, but found the results inconsistent or unhelpful. P12, who works in the financial sector, highlighted the situational use of GenAI tools and shared a preference for AI-based image recognition apps like Seeing AI over asking others for assistance *“For casual tasks like reading currency or simple documents, Seeing AI is my go-to. It’s fast, reliable, and does exactly what I need. I’d still ask my boss before using AI on confidential work documents—because companies can have their own rules.”*

5.6 Scenario 6: Outdoor Spatial Privacy

In this case, many expressed use of GenAI, which are shaped by past training, current usability limitations, and broader social concerns. For instance, P13 shared skepticism toward overreliance on technology. However, she acknowledged growing community interest in tools like GoodMaps⁵, especially as their coverage of indoor spaces expands. Others echoed similar tensions. P12 described how GenAI tools such as Seeing AI felt socially awkward in public spaces: *“You have to hold the phone up as the bus approaches... people can think you’re taking their photos.”*

Limitation & Expectation. Many shared expectations towards ergonomic, hands-free, and privacy-aware design. P11 described the unsolicited and time-consuming experience of being placed in a wheelchair at airports without understanding the range of blindness *“I would really prefer privacy-preserving GenAI that guides without requiring public disclosures and unsolicited wheelchair placement in airports. I can see an option with GenAI to support step-by-step instructions or a human hand-off?”* P9 shared concerns about location privacy *“I don’t want people to know where I am travelling. I have envisioned and meta glass, but I am not sure if they process GPS and camera data locally or if they connect to other apps externally without my consent.”*

⁵<https://goodmaps.com/>

6 Discussion

We presented one of the first investigations into current practices and design opportunities of GenAI for visual privacy among people with visual impairments. Below, we synthesize our findings to show how GenAI-mediated visual privacy is shaped by personal agency, social norms, and technological constraints.

6.1 Notion of Visual Privacy within GenAI.

Our study surfaced two interrelated dimensions of visual privacy for people with visual impairments when interacting with generative AI tools: (1) impression management on how individuals control the way others perceive their appearance and environment, and (2) accountability in handling others’ private visual content.

Our participants expressed mixed feelings about relying on GenAI, similar to prior research on crowd-powered accessibility tools like VizWiz [25]. While our findings indicate the convenience of the GenAI tool in addressing long-standing challenges in managing one’s privacy independently, the impersonal nature and lack of contextual awareness of GenAI could potentially amplify privacy concerns. Participants often use GenAI for socially sensitive tasks that directly shape how others perceive them and their surroundings, or convey specific messages (e.g., competence, professionalism), even for curating the visual backdrop during video calls. These tasks carry implications for social signaling from Goffman’s theory of impression management [88]. However, the generic, de-contextualized feedback from GenAI systems often failed to meet users’ nuanced expectations. For instance, our findings highlight that privacy management varies across different scenarios, while in one case, risks are centered on the user’s own visual contents, in another, risks may arise from shared environments or unintended bystanders. These distinctions underscore the importance of context and relational expectations (e.g., Nissenbaum’s contextual integrity [69]).

Moreover, our findings point to systemic ambiguities in the institutional handling of GenAI use. Participants, particularly blind professionals, described being left to self-govern the use of GenAI tools for high-stakes tasks such as processing confidential disability paperwork, verifying identification, or interpreting student documents. Despite recognizing both the benefits and risks, the lack of formal guidance, secure infrastructure, or vetted tools leads users to make complex privacy decisions independently. This reflects what Ahmed et al. [10] and Akter et al. [14] have identified as structural vulnerabilities, where disabled users shoulder disproportionate responsibility in mediating privacy risks in the absence of institutional safeguards.

6.2 Design Gaps: Privacy-Preserving GenAI Ecosystems

While prior research has examined visual privacy for BLV individuals, particularly through human-mediated services like Aira and Be My Eyes [16, 85], our study extends this work by shedding light on how PVI users re-purpose GenAI tools for both accessibility and privacy-sensitive tasks. While existing research has explored how blind individuals use visual interpretation tools like Aira, Be My Eyes, and Seeing AI, and some emerging tools [7], we highlighted the design gap of GenAI tools for visual privacy management.

For example, participants preferred GenAI over human assistance for sensitive tasks like reading pregnancy tests or mammograms, yet also expressed concern over data storage [77] and potential model training risks [105]. These practices highlight that visual privacy involves not just what is shared, but how and with whom, introducing interpersonal and emotional dimensions of risk. In managing spatial privacy, participants emphasized complementing, rather than replacing, tactile methods with GenAI. This underscores a hybrid design space not yet central in GenAI discussions. Similarly, our findings add to research on online sharing and impression management [24, 70] by surfacing the challenge of distinguishing between one's own and others' belongings in shared environment—an area yet to be integrated in mainstream GenAI tools.

We suggest GenAI design should move beyond basic image description toward adaptive, context-aware, and multimodal feedback that reflects users' privacy priorities. This includes features like personalized object filters and proactive alerts for sensitive content grounded in the real practices and concerns of BLV users.

6.3 Design implications

Drawing upon BLV people's perceived limitations and expectations of GenAI, we suggest actionable design implications.

On-Device or Local Encryption. Our findings revealed a common expectation across all scenarios for local, on-device data processing, particularly for facial images, home environments, financial or medical documents, and GPS or camera inputs used by BLV individuals in both public and private space. This expectation also extended to cases where BLV users processed documents on behalf of others in professional roles (e.g., student papers, social security benefit documents for eligibility). One potential approach is to design a modular, plug-and-play GenAI model sandbox per data type: Visual, Text, Audio/Video, and Sensor Module (GPS) isolated in a trusted execution environment (e.g., Intel TDX, iOS Secure Enclave⁶). This technical design could be paired with audit log and explainability layers to provide privacy logs detailing accessed data, module use, and outputs. This will ensure there is no network access for the sandbox unless explicitly approved by the user.

Federated Compliance-Aware GenAI: Secure Toolkit for BLV Professional. Our findings revealed a distinct concern around protecting others' privacy when BLV professionals use GenAI in institutional contexts like banking, education, or social services, where users often process sensitive data on behalf of others. In response, we propose a secure GenAI toolkit that integrates with compliance-aware platforms (e.g., SharePoint, OneDrive, Google Workspace) to handle documents locally or in secure federated and decentralized environments [64, 78]. It can also enforce role-based access controls (RBAC) and aligns with institutional data protection standards, such as FERPA and HIPAA, through privacy-policy-aware schema. To further safeguard privacy, the system includes post-task accountability prompts using named entity recognition (NER) to flag sensitive data (e.g., names, birthdates) and guide protective actions.

Personalized, Privacy-Aware Appearance Feedback System. Our findings show interest of using of GenAI for self-appearance

management, yet this raises privacy concerns and the need to distinguish bodily privacy (e.g., exposed undergarments) from aesthetic issues (e.g., smudged makeup). To support BLV users, we propose integrating appearance-specific descriptors through a fine-tuned multi-label visual classifier and customizable sensitivity profiles. the system can also feature memory-based look validation, which compares the user's current appearance with previously approved outfits using CLIP-based [71] visual similarity and securely stored, user-labeled photos.

Visual Disambiguation for GenAI for Shared Space Privacy. To support blind users in managing visual privacy in shared environments, GenAI tools should incorporate personalized object recognition and context-aware disambiguation. A technical solution involves allowing users to train the system to identify and differentiate their personal items from those of others using labeled image examples. For instance, when analyzing images, the system can prompt: *"Multiple items detected, do you want to identify which are yours?"* and apply role-specific visual tagging (e.g., not yours, shared, uncertain).

7 Conclusion

Our paper presents one of the first in-depth empirical studies on how blind and low vision (BLV) individuals use Generative AI (GenAI) tools to manage visual privacy across diverse everyday contexts. Our findings extend the existing literature on both privacy and accessibility by demonstrating that visual privacy for BLV users is not just about limiting data disclosure, but also about facilitating autonomy, trust, and social accountability. As GenAI continues to evolve, we call on researchers and designers to prioritize the lived realities of marginalized users not as edge cases, but as starting points for innovation.

Acknowledgments

We thank our participants for their thoughtful insights. This work was partially supported by the National Science Foundation SaTC grants (#2125925, #2148080, #2126314).

References

- [1] [n.d.]. Envision - Perceive Possibility — letsenvision.com. <https://www.letsenvision.com/>. [Accessed 08-04-2025].
- [2] [n.d.]. Home — bemyeyes.com. <https://www.bemyeyes.com/>. [Accessed 08-04-2025].
- [3] [n.d.]. Seeing AI - Talking Camera for the Blind — seeingai.com. <https://www.seeingai.com/>. [Accessed 08-04-2025].
- [4] 2023. Be My Eyes. <https://www.bemyeyes.com/>. Accessed: 2023-9-02.
- [5] Alessandro Acquisti and Jens Grossklags. 2005. Privacy and rationality in individual decision making. *IEEE security & privacy* 3, 1 (2005), 26–33.
- [6] Dustin Adams, Lourdes Morales, and Sri Kurniawan. 2013. A Qualitative Study to Support a Blind Photography Mobile Application. In *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments* (Rhodes, Greece) (PETRA '13). Association for Computing Machinery, New York, NY, USA, Article 25, 8 pages. doi:10.1145/2504335.2504360
- [7] 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*. 1–14.
- [8] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy concerns and behaviors of people with visual impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 3523–3532.
- [9] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy Concerns and Behaviors of People with Visual Impairments. In

⁶<https://docs.trustauthority.intel.com/main/articles/concept-tees-overview.html>

- Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 3523–3532. doi:10.1145/2702123.2702334
- [10] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2016. Addressing physical safety, security, and privacy for people with visual impairments. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. 341–354.
 - [11] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2016. Addressing Physical Safety, Security, and Privacy for People with Visual Impairments. In *Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security* (Denver, CO, USA) (SOUPS '16). USENIX Association, USA, 341–354.
 - [12] Aira. 2020. Aira. <https://aira.io/>
 - [13] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Manohar Swaminathan. 2022. Shared privacy concerns of the visually impaired and sighted bystanders with camera-based assistive technologies. *ACM Transactions on Accessible Computing (TACCESS)* 15, 2 (2022), 1–33.
 - [14] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Manohar Swaminathan. 2022. Shared Privacy Concerns of the Visually Impaired and Sighted Bystanders with Camera-Based Assistive Technologies. *ACM Trans. Access. Comput.* 15, 2, Article 11 (may 2022), 33 pages. doi:10.1145/3506857
 - [15] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. "I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. In *29th USENIX Security Symposium (USENIX Security 20)*. 1929–1948.
 - [16] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. "I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, 1929–1948. <https://www.usenix.org/conference/usenixsecurity20/presentation/akter>
 - [17] Jean-Baptiste Alayrac, Jeff Donahue, Pauline Luc, Antoine Miech, Iain Barr, Yana Hasson, Karel Lenc, Arthur Mensch, Katherine Millican, Malcolm Reynolds, et al. 2022. Flamingo: a visual language model for few-shot learning. *Advances in neural information processing systems* 35 (2022), 23716–23736.
 - [18] Gus Alexiou. [n.d.]. Envision Adds ChatGPT AI Sight Assistance To Its Smart Glasses For The Blind — forbes.com. <https://www.forbes.com/sites/gusalexio/2023/04/30/envision-adds-chatgpt-ai-sight-assistance-to-its-smart-glasses-for-the-blind/?sh=d443c835446f>. [Accessed 08-04-2025].
 - [19] Rahaf Alharbi, Robin N. Brewer, and Sarita Schoenebeck. 2022. Understanding Emerging Obfuscation Technologies in Visual Description Services for Blind and Low Vision People. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 469 (nov 2022), 33 pages. doi:10.1145/3555570
 - [20] Maryam Bandukda, Catherine Holloway, Aneesha Singh, and Nadia Berthouze. 2020. PLACES: a framework for supporting blind and partially sighted people in outdoor leisure activities. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–13.
 - [21] Hanoona Bangalath, Muhammad Maaz, Muhammad Uzair Khattak, Salman H Khan, and Fahad Shahbaz Khan. 2022. Bridging the gap between object and image-level representations for open-vocabulary detection. *Advances in Neural Information Processing Systems* 35 (2022), 33781–33794.
 - [22] Andy Beatman. 2024. 6 ways generative AI helps improve accessibility for all with Azure. <https://azure.microsoft.com/en-us/blog/6-ways-generative-ai-helps-improve-accessibility-for-all-with-azure/>. [Accessed 08-04-2025].
 - [23] Oliver Bendel. 2024. How Can Generative AI Enhance the Well-being of Blind?. In *Proceedings of the AAAI Symposium Series*, Vol. 3. 340–347.
 - [24] Cynthia L. Bennett, Jane E. Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3173574.3173650
 - [25] Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samuel White, et al. 2010. Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology*. 333–342.
 - [26] Jeffrey P. Bigham, Tom Yeh, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Aubrey Tatarowicz, Brandyn White, and Samuel White. 2010. VizWiz: nearly real-time answers to visual questions. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A) - W4A '10*. ACM Press, Raleigh, North Carolina, 1. doi:10.1145/1805986.1806020
 - [27] Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, and Jeffrey P Bigham. 2013. Visual challenges in the everyday lives of blind people. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2117–2126.
 - [28] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association.
 - [29] Finn Brunton and Helen Nissenbaum. 2015. *Obfuscation: A User's Guide for Privacy and Protest*. The MIT Press.
 - [30] Ruei-Che Chang, Yuxuan Liu, and Anhong Guo. 2024. WorldScribe: Towards Context-Aware Live Visual Descriptions. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–18.
 - [31] Chaoran Chen, Daodao Zhou, Yanfang Ye, Toby Jia-jun Li, and Yaxing Yao. 2025. CLEAR: Towards Contextual LLM-Empowered Privacy Policy Analysis and Risk Generation for Large Language Model Applications. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 277–297.
 - [32] Tiffany Chen, Cassandra Lee, Jessica R Mindel, Neska Elhouij, and Rosalind Picard. 2023. Closer worlds: using generative AI to facilitate intimate conversations. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.
 - [33] Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E Gonzalez, et al. 2023. Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. See <https://vicuna.lmsys.org> (accessed 14 April 2023) 2, 3 (2023), 6.
 - [34] Chun Jie Chong, Chenxi Hou, Zhihao Yao, and Seyed Mohammadjavad Seyed Talebi. 2024. Casper: Prompt Sanitization for Protecting User Privacy in Web-Based Large Language Models. *arXiv preprint arXiv:2408.07004* (2024).
 - [35] 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 2024 CHI Conference on Human Factors in Computing Systems*. 1–21.
 - [36] Agebson Rocha Façanha, Ticianne Darin, Windson Viana, and Jaime Sánchez. 2020. O&M indoor virtual environments for people who are blind: A systematic literature review. *ACM Transactions on Accessible Computing (TACCESS)* 13, 2 (2020), 1–42.
 - [37] Yuanyuan Feng, Abhilasha Ravichander, Yaxing Yao, Shikun Zhang, and Rex Chen. 2024. Understanding How to Inform Blind and {Low-Vision} Users about Data Privacy through Privacy Question Answering Assistants. In *33rd USENIX Security Symposium (USENIX Security 24)*. 2065–2082.
 - [38] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods* 5, 1 (2006), 80–92.
 - [39] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. 2013. *Statistical methods for rates and proportions*. John Wiley & sons.
 - [40] Claudia Flores-Saviaga, Benjamin V Hanrahan, Kashif Imteyaz, Steven Clarke, and Saiph Savage. 2025. The Impact of Generative AI Coding Assistants on Developers Who Are Visually Impaired. *arXiv preprint arXiv:2503.16491* (2025).
 - [41] Ricardo Gonzalez, Jazmin Collins, Shiri Azenkot, and Cynthia Bennett. 2024. Investigating Use Cases of AI-Powered Scene Description Applications for Blind and Low Vision People. *arXiv preprint arXiv:2403.15604* (2024).
 - [42] Ricardo E. Gonzalez Penuela, Paul Vermette, Zihan Yan, Cheng Zhang, Keith Vertanen, and Shiri Azenkot. 2022. Understanding How People with Visual Impairments Take Selfies: Experiences and Challenges. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) (ASSETS '22). Association for Computing Machinery, New York, NY, USA, Article 63, 4 pages. doi:10.1145/3517428.3550372
 - [43] Darren Guinness, Edward Cutrell, and Meredith Ringel Morris. 2018. Caption crawler: Enabling reusable alternative text descriptions using reverse image search. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–11.
 - [44] Danna Gurari, Qing Li, Chi Lin, Yanan Zhao, Anhong Guo, Abigale Stangl, and Jeffrey P. Bigham. 2019. VizWiz-Priv: A Dataset for Recognizing the Presence and Purpose of Private Visual Information in Images Taken by Blind People. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
 - [45] Danna Gurari, Qing Li, Abigale J Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. 2018. Vizwiz grand challenge: Answering visual questions from blind people. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3608–3617.
 - [46] Yu Hao, Alexey Magay, Hao Huang, Shuaihang Yuan, Congcong Wen, and Yi Fang. 2024. ChatMap: A Wearable Platform Based on the Multi-modal Foundation Model to Augment Spatial Cognition for People with Blindness and Low Vision. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 129–134.
 - [47] Susumu Harada, Daisuke Sato, Dustin W. Adams, Sri Kurniawan, Hironobu Takagi, and Chieko Asakawa. 2013. Accessible Photo Album: Enhancing the Photo Sharing Experience for People with Visual Impairment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). Association for Computing Machinery, New York, NY, USA, 2127–2136. doi:10.1145/2470654.2481292
 - [48] Jordan Hayes, Smirity Kaushik, Charlotte Emily Price, and Yang Wang. 2019. Cooperative privacy and security: Learning from people with visual impairments and their allies. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 1–20.

- [49] Orit Hazzan and Yael Erez. 2024. Generative AI in computer science education. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 2*. 1899–1899.
- [50] Mina Huh, Yi-Hao Peng, and Amy Pavel. 2023. GenAssist: Making image generation accessible. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–17.
- [51] 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.
- [52] Alekhya Jonnalagedda, Lucy Pei, Suryansh Saxena, Ming Wu, Byung-Cheol Min, Ermine A Teves, Aaron Steinfeld, and M Bernardine Dias. 2014. Enhancing the safety of visually impaired travelers in and around transit stations. *Journal Contribution* (2014), 1–31.
- [53] Smirity Kaushik, Natā M Barbosa, Yaman Yu, Tanusree Sharma, Zachary Kilhofer, JooYoung Seo, Sauvik Das, and Yang Wang. 2023. {GuardLens}: Supporting safer online browsing for people with visual impairments. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. 361–380.
- [54] Majeed Kazemitabaar, Justin Chow, Carl Ka To Ma, Barbara J Ericson, David Weintrop, and Tovi Grossman. 2023. Studying the effect of AI code generators on supporting novice learners in introductory programming. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–23.
- [55] Seonghee Lee, Maho Kohga, Steve Landau, Sile O'Modhrain, and Hari Subramonyam. 2024. AltCanvas: A Tile-Based Editor for Visual Content Creation with Generative AI for Blind or Visually Impaired People. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–22.
- [56] Sooyeon Lee, Rui Yu, Jingyi Xie, Syed Masum Billah, and John M Carroll. 2022. Opportunities for human-AI collaboration in remote sighted assistance. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*. 63–78.
- [57] Franklin Mingzhe Li, Francesca Spektor, Meng Xia, Mina Huh, Peter Cederberg, Yui Gong, Kristen Shinohara, and Patrick Carrington. 2022. "It feels like taking a gamble": Exploring perceptions, practices, and challenges of using makeup and cosmetics for people with visual impairments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [58] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. 2023. Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models. In *International conference on machine learning*. PMLR, 19730–19742.
- [59] KunChang Li, Yanan He, Yi Wang, Yizhuo Li, Wenhui Wang, Ping Luo, Yali Wang, Limin Wang, and Yu Qiao. 2023. Videochat: Chat-centric video understanding. *arXiv preprint arXiv:2305.06355* (2023).
- [60] Feng Liang, Bichen Wu, Xiaoliang Dai, Kunpeng Li, Yanan Zhao, Hang Zhang, Peizhao Zhang, Peter Vajda, and Diana Marculescu. 2023. Open-vocabulary semantic segmentation with mask-adapted clip. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7061–7070.
- [61] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. 2024. Visual instruction tuning. *Advances in neural information processing systems* 36 (2024).
- [62] Ying Ma, Shiquan Zhang, Dongju Yang, Zhanna Sarsenbayeva, Jarrod Knibbe, and Jorge Goncalves. 2025. Raising Awareness of Location Information Vulnerabilities in Social Media Photos using LLMs. *arXiv preprint arXiv:2503.20226* (2025).
- [63] Muhammad Maaz, Hanoona Rasheed, Salman Khan, and Fahad Shahbaz Khan. 2023. Video-chatgpt: Towards detailed video understanding via large vision and language models. *arXiv preprint arXiv:2306.05424* (2023).
- [64] Philippe Massonet, Syed Naqvi, Christophe Ponsard, Joseph Latanicki, Benny Rochwerger, and Massimo Villari. 2011. A monitoring and audit logging architecture for data location compliance in federated cloud infrastructures. In *2011 IEEE international symposium on parallel and distributed processing workshops and PhD forum*. IEEE, 1510–1517.
- [65] Meredith Ringel Morris, Annuska Zolyomi, Catherine Yao, Sina Bahram, Jeffrey P Bigham, and Shaun K Kane. 2016. "With most of it being pictures now, I rarely use it" Understanding Twitter's Evolving Accessibility to Blind Users. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5506–5516.
- [66] National Federation of the Blind. [n. d.]. Blind Users Innovating and Leading Design. <https://nfb.org/programs-services/center-excellence-nonvisual-access/blind-users-innovating-and-leading-design>
- [67] Bolin Ni, Houwen Peng, Minghao Chen, Songyang Zhang, Gaofeng Meng, Jianlong Fu, Shiming Xiang, and Haibin Ling. 2022. Expanding language-image pretrained models for general video recognition. In *European Conference on Computer Vision*. Springer, 1–18.
- [68] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
- [69] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.
- [70] Soobin Park. 2020. Supporting Selfie Editing Experiences for People with Visual Impairments. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, Greece) (ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 106, 3 pages. doi:10.1145/3373625.3417082
- [71] Yuxin Peng and Chong-Wah Ngo. 2006. Clip-based similarity measure for query-dependent clip retrieval and video summarization. *IEEE Transactions on Circuits and Systems for Video Technology* 16, 5 (2006), 612–627.
- [72] Ricardo E Gonzalez Penuela, Ruiying Hu, Sharon Lin, Tanisha Shende, and Shiri Azenkot. 2025. Towards Understanding the Use of MLLM-Enabled Applications for Visual Interpretation by Blind and Low Vision People. *arXiv preprint arXiv:2503.05899* (2025).
- [73] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. 2021. Learning transferable visual models from natural language supervision. In *International conference on machine learning*. PMLR, 8748–8763.
- [74] Gayatri Raman and Erin Brady. 2024. Exploring Use and Perceptions of Generative AI Art Tools by Blind Artists. *arXiv preprint arXiv:2409.08226* (2024).
- [75] Hanoona Rasheed, Muhammad Uzair Khattak, Muhammad Maaz, Salman Khan, and Fahad Shahbaz Khan. 2023. Fine-tuned clip models are efficient video learners. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6545–6554.
- [76] David Rozenberszki, Or Litany, and Angela Dai. 2022. Language-grounded indoor 3d semantic segmentation in the wild. In *European Conference on Computer Vision*. Springer, 125–141.
- [77] Tanusree Sharma, Lin Kyi, Yang Wang, and Asia J Biega. 2024. "I'm not convinced that they don't collect more than is necessary": {User-Controlled} Data Minimization Design in Search Engines. In *33rd USENIX Security Symposium (USENIX Security 24)*. 2797–2812.
- [78] Tanusree Sharma, Vivek C Nair, Henry Wang, Yang Wang, and Dawn Song. 2024. "I Can't Believe It's Not Custodial!": Usable Trustless Decentralized Key Management. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [79] Tanusree Sharma, Abigale Stangl, Lotus Zhang, Yu-Yun Tseng, Inan Xu, Leah Findlater, Danna Gurari, and Yang Wang. 2023. Disability-first design and creation of a dataset showing private visual information collected with people who are blind. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [80] Fangxun Shu, Lei Zhang, Hao Jiang, and Cihang Xie. 2023. Audio-visual llm for video understanding. *arXiv preprint arXiv:2312.06720* (2023).
- [81] Abigale Stangl, Meredith Ringel Morris, and Danna Gurari. 2020. "Person, Shoes, Tree. Is the Person Naked?" What People with Vision Impairments Want in Image Descriptions. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.
- [82] Abigale Stangl, Emma Sadjo, Pardis Emami-Naeini, Yang Wang, Danna Gurari, and Leah Findlater. 2023. "Dump It, Destroy It, Send It to Data Heaven": Blind People's Expectations for Visual Privacy in Visual Assistance Technologies. In *Proceedings of the 20th International Web for All Conference (Austin, TX, USA) (W4A '23)*. Association for Computing Machinery, New York, NY, USA, 134–147. doi:10.1145/3587281.3587296
- [83] Abigale Stangl, Emma Sadjo, Pardis Emami-Naeini, Yang Wang, Danna Gurari, and Leah Findlater. 2023. "Dump it, Destroy it, Send it to Data Heaven": Blind People's Expectations for Visual Privacy in Visual Assistance Technologies. In *Proceedings of the 20th International Web for All Conference*. 134–147.
- [84] Abigale Stangl, Kristina Shiroma, Bo Xie, Kenneth R Fleischmann, and Danna Gurari. 2020. Visual content considered private by people who are blind. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–12.
- [85] Abigale Stangl, Kristina Shiroma, Bo Xie, Kenneth R. Fleischmann, and Danna Gurari. 2020. Visual Content Considered Private by People Who Are Blind. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, Greece) (ASSETS '20)*. Association for Computing Machinery, New York, NY, USA, Article 31, 12 pages. doi:10.1145/3373625.3417014
- [86] Bolun Sun, Yifan Zhou, and Haiyun Jiang. 2024. Empowering Users in Digital Privacy Management through Interactive LLM-Based Agents. *arXiv preprint arXiv:2410.11906* (2024).
- [87] Issatay Tokmurziyev, Miguel Altamirano Cabrera, Muhammad Haris Khan, Yara Mahmoud, Luis Moreno, and Dzmitry Tsetserukou. 2025. LLM-Glasses: GenAI-driven Glasses with Haptic Feedback for Navigation of Visually Impaired People. *arXiv preprint arXiv:2503.16475* (2025).
- [88] Efrat Tseñlon. 1992. Is the presented self sincere? Goffman, impression management and the postmodern self. *Theory, culture & society* 9, 2 (1992), 115–128.
- [89] Yu-Yun Tseng, Tanusree Sharma, Lotus Zhang, Abigale Stangl, Leah Findlater, Yang Wang, and Danna Gurari. 2024. BIV-Priv-Seg: Locating Private Content in Images Taken by People With Visual Impairments. *arXiv preprint arXiv:2407.18243* (2024).

- [90] Sheena Vasani. [n. d.]. Be My Eyes AI offers GPT-4-powered support for blind Microsoft customers — theverge.com. <https://www.theverge.com/2023/11/15/23962709/microsoft-blind-users-open-ai-chatgpt-4-be-my-eyes>. [Accessed 08-04-2025].
- [91] Violeta Voykinska, Shiri Azenkot, Shaomei Wu, and Gilly Leshed. 2016. How Blind People Interact with Visual Content on Social Networking Services. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (San Francisco, California, USA) (CSCW '16). Association for Computing Machinery, New York, NY, USA, 1584–1595. doi:10.1145/2818048.2820013
- [92] Violeta Voykinska, Shiri Azenkot, Shaomei Wu, and Gilly Leshed. 2016. How blind people interact with visual content on social networking services. In *Proceedings of the 19th acm conference on computer-supported cooperative work & social computing*. 1584–1595.
- [93] Herman Wahidin, Jenny Waycott, and Steven Baker. 2018. The challenges in adopting assistive technologies in the workplace for people with visual impairments. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. 432–442.
- [94] Bryan Wang, Yuliang Li, ZhaoYang Lv, Haijun Xia, Yan Xu, and Raj Sodhi. 2024. LAVE: LLM-powered agent assistance and language augmentation for video editing. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 699–714.
- [95] Mengmeng Wang, Jiazheng Xing, and Yong Liu. 2021. Actionclip: A new paradigm for video action recognition. *arXiv preprint arXiv:2109.08472* (2021).
- [96] Qi Wu, Peng Wang, Chunhua Shen, Anthony Dick, and Anton Van Den Hengel. 2016. Ask me anything: Free-form visual question answering based on knowledge from external sources. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4622–4630.
- [97] Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic alt-text: Computer-generated image descriptions for blind users on a social network service. In *proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 1180–1192.
- [98] Jingyi Xie, Madison Reddie, Sooyeon Lee, Syed Masum Billah, Zihan Zhou, Chun-Hua Tsai, and John M Carroll. 2022. Iterative design and prototyping of computer vision mediated remote sighted assistance. *ACM Transactions on Computer-Human Interaction (TOCHI)* 29, 4 (2022), 1–40.
- [99] Jingyi Xie, Rui Yu, He Zhang, Syed Masum Billah, Sooyeon Lee, and John M Carroll. 2025. Beyond Visual Perception: Insights from Smartphone Interaction of Visually Impaired Users with Large Multimodal Models. *arXiv preprint arXiv:2502.16098* (2025).
- [100] Jingyi Xie, Rui Yu, He Zhang, Sooyeon Lee, Syed Masum Billah, and John M Carroll. 2024. Bubblecam: Engaging privacy in remote sighted assistance. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.
- [101] Lotus Zhang, Abigale Stangl, Tanusree Sharma, Yu-Yun Tseng, Inan Xu, Danna Gurari, Yang Wang, and Leah Findlater. 2024. Designing Accessible Obfuscation Support for Blind Individuals' Visual Privacy Management. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [102] Shuning Zhang, Xin Yi, Haobin Xing, Lyumanshan Ye, Yongquan Hu, and Hewu Li. 2024. Adanonymizer: Interactively Navigating and Balancing the Duality of Privacy and Output Performance in Human-LLM Interaction. *arXiv preprint arXiv:2410.15044* (2024).
- [103] Xinyu Zhang, Huiyu Xu, Zhongjie Ba, Zhibo Wang, Yuan Hong, Jian Liu, Zhan Qin, and Kui Ren. 2024. Privacyasst: Safeguarding user privacy in tool-using large language model agents. *IEEE Transactions on Dependable and Secure Computing* (2024).
- [104] Zhuohao (Jerry) Zhang, Smirity Kaushik, JooYoung Seo, Haolin Yuan, Sauvik Das, Leah Findlater, Danna Gurari, Abigale Stangl, and Yang Wang. 2023. ImageAlly: A Human-AI Hybrid Approach to Support Blind People in Detecting and Redacting Private Image Content. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX Association, Anaheim, CA, 417–436. <https://www.usenix.org/conference/soups2023/presentation/zhang>
- [105] Jijie Zhou, Eryue Xu, Yaoyao Wu, and Tianshi Li. 2024. Rescriber: Smaller-LLM-Powered User-Led Data Minimization for Navigating Privacy Trade-offs in LLM-Based Conversational Agent. *arXiv preprint arXiv:2410.11876* (2024).
- [106] Deyao Zhu, Jun Chen, Xiaoqian Shen, Xiang Li, and Mohamed Elhoseiny. 2023. Minigt-4: Enhancing vision-language understanding with advanced large language models. *arXiv preprint arXiv:2304.10592* (2023).

A Appendix

For these informal interviews, I'm only considering people who used GAI before. I am first going to ask you a round of questions so that I can get a better grasp of your current practice and experience with Generative AI Tools. ["Generative AI for Privacy" if any]

A.1 Interview

Section 1: RQ1: How and for what purpose do blind users use Generative AI tools?

What generative AI tools are you currently using?

- ChatGPT
- DALL.E.2
- Be My AI
- Envision AI
- Gemini
- Claude
- Others

Which ones have you used specifically to access visual information, like images, videos, or real-world visual information?

- Be My AI
- Envision AI
- ChatGPT
- Others

For each of the tools you mentioned, what purposes do you use them for? Let's start with one by one.

- Can you share an experience or event when you used this tool (e.g., Envision AI)?
- What makes you decide to use this tool? (Some factors: they might share some insights on its usefulness, performance, new features, etc.)
- Could you briefly explain how this tool works? Maybe you could provide an example and walk me through your process of using this tool.
- Great, have you found these features in any existing tools you previously used? What is different about generative AI compared to past tools you used for this similar purpose?
- It's great to hear about how and for what purpose you use this tool. Is there any other purpose you consider using this tool for? If not, can you think of any other way this tool could be useful in your daily tasks?

Based on their response, if no privacy-related use cases for Generative AI tools are mentioned, then proactively ask about privacy-related tasks for each tool they mentioned:

- Have you used these tools for any purpose related to privacy? (Open-ended for them to answer)

Definition of the Scope: Visual Privacy for Blind Users refers to the safeguarding and management of sensitive visual information that could be shared or disclosed through the use of Generative AI tools. This includes but is not limited to the protection of content such as medical records, financial information, or any other visual data that might be considered private when engaging in daily activities. For instance, when using these tools to receive descriptions of potentially sensitive content—whether it's medical or financial records or when scanning surroundings for navigation.

Next, I will ask you about specific scenarios to better understand how you currently use Generative AI tools and to explore your thoughts on how you might want to use them in similar situations in the future.

Section 2: Future Design Preferences for Generative AI Tools

Privacy Scenario 1 (Own Privacy): Checking Self Before Seeing Other People Imagine you are about to meet someone at a social event or professional meeting. Before you step out, you want to make sure your appearance is in order—clothes are neat, hair is tidy, and nothing is out of place. You decide to use a Generative AI tool to describe your appearance.

- How would you currently use a Generative AI tool to check your appearance before meeting someone?
- How would you like this tool to be designed in the future to better assist you in this scenario?
- What is it about GAI that makes you think this way (risk perception)?

Privacy Scenario 2 (Own Privacy): Checking House Before Having Someone Over Imagine you're expecting guests at your home and want to ensure that your space is tidy and presentable. You want to check for things like clutter in the living room, any dishes left out in the kitchen, or if there's anything unusual that might be out of place. You decide to use a Generative AI tool to help you assess your surroundings.

- How would you currently use a Generative AI tool to check your house before having someone over or in some similar situation?
- How would you like this tool to be designed in the future to better assist you in this scenario?

Privacy Scenario 3 (Own Privacy): Picture to Share on Social Media Imagine you've taken a photo and are considering sharing it on social media. Before posting, you want to ensure that the image doesn't contain any private information, such as personal things, recognizable locations, or other details that you wouldn't want to be publicly visible. You decide to use a Generative AI tool to analyze and describe its contents.

- How would you currently use a Generative AI tool to check the content of a picture before sharing it on social media?
- How would you like this tool to be designed in the future to better assist you in this scenario?

Privacy Scenario 4 (Own Privacy): Document Picture to Share With Employer Imagine you're preparing to share a financial report with an employee, but before doing so, you want to ensure that the document doesn't contain personal salary details, confidential company financials, or other data that you don't want to disclose.

- How would you currently use a Generative AI tool to check such a document before sharing it with an employee?
- How would you like this tool to be designed in the future to better assist you in this scenario?

Privacy Scenario 5 (Others' Privacy): Reading Private Documents (Social Security) for Issuing Benefits for Employees As an employer, you are responsible for issuing benefits to your employees, which requires you to review private documents, such as Social Security information, tax forms, or other sensitive personal data. You decide to rely on a Generative AI tool to assist you in reading this information while ensuring that privacy and confidentiality are maintained.

- How would you currently use a Generative AI tool to read and process documents of your employees in such a scenario for issuing employee benefits?
- How would you like this tool to be designed in the future to better assist you in this scenario?

Privacy Scenario 6 (Others' Privacy): Scanning Surroundings When Outside (Navigating the Airport) Imagine you are at an airport and need to navigate through various areas, such as finding your gate, locating restrooms, or identifying nearby amenities. You decide to rely on a Generative AI tool to help you understand and navigate the environment.

- How would you currently use a Generative AI tool to scan and understand your surroundings in an airport?
- How would you like this tool to be designed in the future to better assist you in this scenario?

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009