"Dump it, Destroy it, Send it to Data Heaven": Blind People's Expectations for Visual Privacy in Visual Assistance Technologies

Abigale Stangl
University of Washington
Seattle, Washington, USA
astangl@uw.edu

Emma Sadjo
University of Washington
Seattle, Washington, USA
esadjo@uw.edu

Pardis Emami-Naeini
University of Washington
Seattle, Washington, USA
pardis@cs.washington.edu

Yang Wang
University of Illinois
Urbana, Illinois, USA
yvw@illinois.edu

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

Leah Findlater
University of Washington
Seattle, Washington, USA
leahkf@uw.edu

ABSTRACT

Visual assistance technologies provide people who are blind with access to information about their visual surroundings by digitally connecting them to remote humans or artificial intelligence systems that describe visual content such as objects, people, scenes, and text observed in their live image/video feeds. Prior work has revealed that users have concerns about how such technologies handle private visual content captured in their image/video feeds. Yet, it remains unclear how users want technologies to manage such private content. To fill this gap, we interviewed 16 totally blind individuals to learn about their expectations for visual privacy when using visual assistance technologies. Our findings reveal three overarching user-centered expectations associated with visual privacy-preservation in this domain, as well as the broader ethical challenges involved with developing AI-based privacy-preserving visual assistance technologies.

ACM Reference Format:
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 *20th International Web for All Conference (W4A '23), April 30–May 01, 2023, Austin, TX, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3587281.3587296

1 INTRODUCTION

Visual assistance technologies (VATs) are used by hundreds of *thousands* of individuals who are blind around the world [106] to accomplish everyday tasks such as navigating, shopping, and completing forms. With the VAT application Seeing AI [88], an individual can use their smartphone to capture an image of a letter they have received and hear the address and content of the letter read aloud. As another commercial example, with Aira [9], a blind person can use their phone to scan an unfamiliar area and stream the video over the internet to a remote-sighted assistant who describes what they see. Over the past 15 years, the marketplace for VATs has boomed, partly due to the development of artificial intelligence (AI) to augment the human labor traditionally needed to provide visual assistance [106].

VATs provide an important assistive service for blind people to access information about their visual surroundings, but users of VATs also encounter a privacy trade-off [113]. As shown in one study, approximately 10% of roughly 40,000 photos that individuals who are blind shared with a VAT contained *private visual content*, such as medical, financial, proprietary, and biometric content [6, 8, 54, 113]. Blind users experience privacy concerns[1] for themselves and for bystanders captured in their images/videos [5, 7, 11–13, 59, 113].

Users commonly feel uninformed about how VATs handle their visual data [112], indicating a mismatch between their need for information and what VATs provide [112]. In 2019-2021, VAT companies rarely provided notice about the handling of images/videos in their privacy policies [112]. When mismatches occur, unpleasant surprises, discomforts, and concerns arise for users [36, 56, 79, 103]. In contrast, when technologies behave as **expected** by the user, fewer privacy concerns arise. Rao et al. [100] identifies that *desires, predictions, rights* and *tolerances* form expectations.

In this paper, we present user-centered research to identify *expectations that blind individuals have for VATs to preserve their visual privacy* during assistance sessions and other stages of the data lifecycle. We share insight on how technology companies can preserve users' privacy *in alignment* with their users' expectations. Though the topic of privacy expectations is studied for other technology (e.g., [37, 79, 100]), this has not yet been examined for VATs.

Capturing visual privacy expectations and incorporating them into the design of VATs should enhance the usability and adoption of these technologies. Notably, we emphasize the nature of *visual* privacy, rather than privacy more generally. This is in trend with other contemporary research, which makes this distinction due to the increased pervasiveness of cameras and the growing need for privacy-enhancement of visual media (e.g., [33, 97, 109, 120]). We address the following research questions (RQ): **RQ (1)** *What are blind people's expectations for privacy-preservation when using visual*

W4A '23, April 30–May 01, 2023, Austin, TX, USA

ACM ISBN 979-8-4007-0748-3/23/04.
https://doi.org/10.1145/3587281.3587296

[1] Concerns entail people's affective response to an event or related action, as well as their perception of potential negative consequences associated with sharing information [17]. Concerns also constitute the first step in one's appraisal and management of their experience and response to a socio-technical system.

assistance technologies?; **RQ (2)** *What are blind people's reactions to potential novel techniques that could be used for visual privacy-preservation?*

To address this question, we conducted 1.5-hour semi-structured interviews with 16 people who self-identified as totally blind. The interviews solicited reflections about participants' experiences using VATs, their consideration of visual privacy, and their expectations of VATs to provide privacy-preserving features. We inductively and deductively analyzed the interview data using thematic analysis [26] and guided by Rao's model of privacy expectations [100].

Our findings reveal three overarching user-centered expectations for VATs based on interviews with 16 blind VAT users: (1) never collect images/videos that contain private visual content; (2) develop interactive features that provide logs, in real-time, of who handles their visual data and why, while providing non-disclosure agreements from software developers and third parties; (3) practice caution and ethical consideration during the development and deployment of privacy-preserving AI by including individuals who are blind as co-developers. We further discuss the reality of participants' expectations for AI as well as the ethical considerations that make visual privacy-preservation challenging in this domain.

2 BACKGROUND

Our research is at the intersection of visual assistance technologies (VATs), visual privacy as experienced by individuals who are blind, and usable privacy-preserving technologies.

2.1 Visual Assistance Technologies

With visual assistance technologies (VATs)[2], a user captures an image/video of their surroundings using a camera-based device, such as a smartphone[3] or smart glasses[4], and shares the media with an application created by a VAT company. While many VATs return a description of each shared image/video, some VATs also allow users to include a specific question about the visual content (e.g., [9]). Prior work has largely focused on VATs that use the interpretive skills of paid or volunteer *remote sighted assistants* (e.g., [22, 66, 74, 76]). *Remote sighted assistants* interpret visual information so that the content and purpose of images/videos/live-feed are available non-visually [75], and they may be professionals (e.g., Aira [9]), volunteers (e.g., Be My Eyes [39]), or crowd-workers (e.g. VizWiz, [22]). An increasing number of VATs integrate AI models to automate the generation of image/video descriptions (e.g., [23, 50, 124, 125]), as well as a combination of human and AI [76, 126] to provide visual assistance.

To advance the development of VATs, prior research has explored ways to effectively deliver description and visual question answers (e.g., [47, 48, 71, 91, 98, 114]), as well as the visual information blind people want in the descriptions and when (e.g., [24, 51, 55, 111]). This includes investigations into description authorship by crowdworkers'[110], professionals [75], and AI [19, 71, 104, 122]. Prior work has also investigated different approaches to training professional remote visual assistants (e.g., [75, 90]), as well as approaches that may be used to preserve blind people's visual privacy (e.g., [15, 52]). We extend prior work related to visual privacy and VATs to identify strategies that VATs may take to meet the expectations of VAT users throughout the data life cycle.

[2]Within the blind community, some people refer to VATs as visual interpreters, while others refer to VATs as camera-based assistive technologies, e.g. [12].
[3]VATs that work on mobile devices include: Aira [9], Seeing AI [88], Be My Eyes [39], Envision [1] TapTapSee [117], Be Specular [20], Supersense [115], LookTel Money Reader [102], Orcam [96], and KNFB Reader [101].
[4]VATs that work on smart glasses include: [9], Orcam [96], and eSite [38].

2.2 Visual Privacy in Visual Assistance Technologies

Visual and aural eavesdropping, insecurity of online transactions, multiparty privacy conflicts, and inadequate information management behavior are common privacy concerns experienced by individuals who are blind [6, 8, 11, 25, 25, 74]. Akter et al. (2022) [12] identify the shared privacy concerns between bystanders and blind people who capture them in the background of their images (i.e., multiparty privacy conflicts). More broadly, Stangl et al. [113] identify that blind people's visual privacy concerns can be attributed to three factors: (1) lack of understanding of how the service is provided (e.g., unknown data handling), (2) expected personal or social impacts (e.g., identity or financial theft), and (3) misalignment with underlying values (e.g., control).

Departing from investigations on visual privacy concerns, we investigate blind people's privacy *expectations* in the context of their formative experiences, values, and the data life-cycle. In a study on privacy expectations, Loi et al. [82] explain, *"Individuals have privacy when established expectations regarding the way information should be transmitted are respected—this is compatible with people expecting different people in different contexts to handle their information in very different ways"*. Understanding users' privacy expectations (i.e., desires, predictions, tolerances, and rights) [100]) is helpful for defining the norms around what information to gather and for what purposes [107], as well as for influencing their privacy decisions around how to use a technology [79]. While prior work has used expectations as a construct when studying touchscreen interfaces [61], web accessibility [10], museums [27], daily food preparation [70], and how to overcome (limiting) expectations (e.g, [67]), to our knowledge, we are the first to investigate expectations for visual privacy and VATs.

2.3 Privacy-Preserving Technologies

Christen et al. [31] define five categories of contemporary privacy-preserving technologies, including mechanisms for: identity authentication and anonymity, private communications, privacy-preserving computations, privacy in databases, and discrimination prevention in data mining. In addition, there are areas where policy measures can offer protections (e.g., consumer data handling, information about breaches, threat intelligence, data sharing, and vulnerability identification), as well as usable privacy features designed for users to engage in personal privacy management (e.g., alternatives to privacy policies like privacy nutrition labels [35] or comic-based policies [116], privacy policy analyzers [58, 128] and summarizers [127], privacy centers, privacy on-boarding flows, and privacy dashboards). In this paper, we first ask participants about their expectations for visual privacy-preserving technologies, and in the context of specific recipients. We also inquire about an emerging privacy-preserving technique (i.e., obfuscation), which is related

to *redaction and sanitization of data* traditionally used to provide privacy in databases [31].

Automatic obfuscation involves computational recognition of private visual content, flagging that content, and application of algorithmic techniques that affect the visual appearance of the flagged content. Automated obfuscation techniques include blurring targeted content, pixelation, overlays of dark pixels, and inpainting [78]. Application of obfuscation techniques commonly provides users with access control by offering options to restrict recipients' views of parts of specific image content, or replace the original content with substitute content (e.g., [62]). Most related to our work, Alharbi et al. [15] take a feminist disability perspective to investigate blind people's responses to obfuscation. They identify that blind VAT users predict that obfuscation is useful for gaining control over text, impression management, and enhancing visual description, but also tensions related to automated obfuscation (e.g., misrecognition, obscuration), and the extreme need for choice. Our findings provide additional depth on how obfuscation can be implemented to provide choice through the process of obfuscation.

3 RESEARCH DESIGN AND METHODS

To investigate blind people's expectations of visual privacy when using VATs (RQ1) and their reactions to potential novel techniques that could be used for visual privacy-preservation (RQ2), we conducted semi-structured interviews with 16 totally blind participants using a protocol approved by our Institutional Review Board (IRB).

3.1 Interview Protocol

We designed interview questions to capture participants' experiences using VATs, their familiarity with visual privacy, their familiarity with visual privacy-preserving features, other contextual factors related to privacy, and their expectations for potential novel privacy-preserving techniques. The design of the protocol was guided by Rao's conceptual framework on privacy expectations, which incorporates privacy theories from Altman [16] and Nissenabuam [93]. The protocol also asks about their familiarity with automatic image/video description, privacy-preserving technologies, prior experience working and collaborating with individuals who are blind, and informal evaluations of VATs to become familiar with existing privacy-preserving mechanisms in VATs. The protocol can be found in the Supplementary Materials.

3.2 Data Collection

The research team is composed of members trained and/or working in computer science, human-computer interaction, or human-centered design. Each member shared their expertise in accessibility, usable privacy, interaction design, machine learning, and/or computer vision through team meetings and written feedback. The first author conducted the interviews with the ($N = 16$) participants remotely over Zoom [129] in autumn 2021. The first researcher recorded the session, reviewed the consent document with the participants, led the interviews, and oversaw the transcription process and data management. A second researcher (one of three undergraduate research assistants) took notes and observed. One of the research assistants is the second author. To account for biases at the outset, all authors are fully sighted.

The interviews lasted between 1.5 and 2 hours[5]. After each interview, we immediately de-identified each participant's data using unique identifier numbers (e.g., [P01]). The interviews were then transcribed using either a third-party human transcription service or a combination of an AI tool for initial transcription (Descript [34]), followed by manual correction by a member of the research team. We consolidated all data into a datasheet that we uploaded to a mixed methods analysis tool MAXQDA [85]) using each interview question as the organizing code category.

3.3 Data Analysis

The first author led a thematic analysis of the data with the three research assistants. We followed a 5-step procedure recommended by Braun and Clarke [26] to create a codebook. Here we describe the process of identifying and organizing themes.

3.3.1 Inductive and Deductive. When cleaning and organizing interview transcripts, we *inductively* assigned a word, short phrase, or *In Vivo* quote[6] to a passage of the interview that conveyed the basic topic and wrote analytic memos to capture initial concepts [7]. We also *deductively* analyzed the data once it was cleaned, by organizing them through an iterative process using the information structure provided by MAXQDA [85]. The structure presented in the findings is guided by four initial parent codes: *privacy values-based, data life-cycle, past/future*, and *expectation type*—as guided by the background literature in **Section 2**.

Two parent codes were *values* and *data life-cycle*. *Values* are the fundamental beliefs that guide our attitudes and actions [43]; participants commonly spoke about control[8], trust[9], and accountability/transparency[10], which we assigned as child codes. The parent code *data life-cycle* included child codes of communication of privacy practices, data collection, and storage, data processing, and data security management. Based on the initial protocol design, we also identified instances when participants' responses pertained to the retelling of a *formative past experience* versus an *expectation* rooted in a prediction, desire, tolerance, and/or right [100]). Finally, once all the statements associated with the interview questions were coded accordingly, two researchers iteratively identified subthemes. We present the findings according to this scaffolding based on the nature of the protocol and our desire to present findings so VATs can use them to develop usable privacy-preserving features at different stages of the data life-cycle.

[5] Six of the interviews were briefly paused due to an incoming call on the participants' end, Zoom crashing, or a drop in internet.
[6] Some codes were *In Vivo*, using short phrases from the participants' own language [89].
[7] When applying codes to each excerpt, we reviewed the excerpt within the original transcript if more context was needed.
[8] *Control* is the power to influence personal effects (including visual data) and experiences [105].
[9] *Trust and confidentiality* relate to the belief in the integrity, ability, or character of a person or thing, and one's confidence or reliance in self, other, and society [119].
[10] *Information-use transparency* is the extent to which an online company allows consumers to access the data collected about them and inform them about how and for what purposes the acquired information is used [69].

3.4 Participant Recruitment

We recruited participants through organizations serving individuals who are blind[11] and via social media (e.g., first author's social media profiles). By conducting a screening survey, we looked for participants who satisfied inclusion criteria: participants had to (1) be totally blind from acquired or congenital blindness, (2) use a VAT weekly or more frequently, and (3) be 18+ years old. In total, 143 people completed the screening survey. Most (n=129) reported using Aira, Seeing AI, or Be My Eyes as their primary VAT. The VATs reported by the remaining 14 people were not comparable in feature sets to these three top VATs (e.g., KNFB and Voice Dream Scanner only read text)–except for Envision AI, which offers optical character recognition (OCR) and automated object recognition. Thus, to gather experiences across full-featured VATs, we used stratified random sampling to select 5 people who used Aira, Seeing AI, and Be My Eyes, and included one Envision AI participant. In total, there were 16 participants.

4 FINDINGS

Here we present findings on (RQ 1) blind people's expectations for visual privacy-preservation when using visual assistance technologies (VATs), and (RQ2) participants' reactions to the idea of AI-based privacy-preserving techniques. We premise these findings with a summary of the participant's demographics and formative experiences that influenced their expectations of privacy.

4.1 Demographics and Experiences

The participants ($N = 16$) varied according to their demographics. Nine self-identified as female and seven as male, between 19 and 72 years old ($M = 43.5$, $SD = 15.9$). With respect to their ethnicity, thirteen participants self-identified as 'White', one as 'Asian', one as 'Black', and one as 'Mixed' ethnicity. Participants completed levels of education varied: seven high school diplomas, two had a Bachelor's degree, six had a Master's degree, and one had a Ph.D. Ten participants were employed full-time, five were part-time, and one sought opportunities. Fifteen were in the U.S., distributed across 12 U.S. states, and one resided in the UK. Additionally, three participants were born outside the United States. All participants identified as totally blind, 10 of whom had congenital blindness, and six of whom had acquired blindness between the ages of 6-12 years old.

The type and range of VATs used varied by the participant; six participants reported using Aira the most, four reported Be My Eyes, five reported Seeing AI, and one reported using Envision AI. All participants used VATs weekly to accomplish everyday tasks (e.g., potting plants, selling items on eBay, washing clothes, and reading mail), which aligns with prior work on blind people's use of VATs (e.g., [24, 30, 46–49, 51, 55, 111]). Participants self-reported their *level of experience with their primary VAT app* as intermediate ($N = 3$), advanced ($N = 7$), and expert ($N = 6$)[12]. Assessing their own expertise using VATs, participants' criteria for their self-assessments: the range of features they use, their frequency of use, their ability to configure settings, and their experience training other people to use VATs. Most of the 16 participants rated visual privacy in their daily lives to be either "absolutely essential" (N=8) or "very essential" (N=4), while the remaining four participants deemed it of at least some ("average") importance, based on a 5-point scale[13], and a common definition of visual privacy.[14] In short answer responses explaining their scores, participants attributed the high importance scores to their desire to maintain *control* over their private visual content. Participants discussed their awareness of visual privacy in terms of prior experiences using digital technologies (e.g., a pop-up asking for permission to record a video call), the impacts of interacting with colleagues, assistants, and allies in sensitive situations (e.g., experiencing their own image being taken without consent, or inadvertently capturing other people in their image) and the potential for interpersonal safety risks (e.g., multi-party privacy conflicts, impression management). Notably, on interpersonal safety, [P10] recounted learning about a pregnancy test result through a remote-sighted assistant, who gained their trust when helping confirm the result at a time when they *"didn't want anybody in my life to know that yet until I knew for sure"*. In this case, the remote-sighted assistants alleviated [P10]'s privacy risk by acting as a shield against having to disclose such private information to others in their life. All the while, highlighting that visual privacy is contextual to the person [93], [P07] shared their ambivalence to having their pregnancy tests read by remote sighted assistants: *"With pregnancy tests, I absolutely get a person I know well [friend or family member] to do that"*).

Importantly, participants reported they had not directly experienced privacy violations when using VATs (e.g., *"I don't remember ever feeling, 'Oh no, I shouldn't have shared that.' 'Oh no, they saw this.'"*[P09]) (despite previously reported visual privacy concerns during use of VATs [13, 112]). Tactics they use to avoid visual privacy risks include (1) Requesting close social ties to report and describe potential or unwanted private visual content disclosures in images/videos in real-time, (2) Forming explicit social agreements about when and where images/videos can be captured in their homes, (3) Creating designated areas in their homes so they can take images/videos without private content, and (4) Focusing cameras on the content of interest only. VATs may develop features to support these privacy-preserving practices.

4.2 Users' Expectations for Visual Privacy-Preservation from VATs

While participants did not report direct experiences of having their visual privacy violated when using VATs, they predicted that *other* individuals who are blind disclose their visual privacy inadvertently, commonly referred to as the *third person effect*[15]. For instance, regarding blind people's visual privacy from a high level, [P15] shared,

[11] The National Federation of the Blind approved and distributed the announcement on their Blind Users Innovating and Leading Design (BUILD) mailing list (e.g., [95]).

[12] Scale for their level of experience using VATs: 1 - fundamental awareness (basic knowledge); 2 - novice (limited experience); 3 - intermediate (practical application); 4 - advanced (applied theory); 5 - expert (recognized authority)

[13] Level of visual privacy importance: 0 - not important at all, 1 - of little importance, 2 - of average importance, 3 - very important, 4 - absolutely essential.

[14] Participants rated the importance of visual privacy in their lives based on a definition of visual privacy as *relationship between the collection, dissemination, and use of visual information, the expectation of privacy, and the legal issues surrounding them prior to asking them to rate the importance of visual privacy in their lives.*

[15] *Third person effects* arise when the primary user believes they will not encounter negative privacy experiences, but other people are likely to have their privacy violated. This effect shows a biased or optimistic perception that can lead to people's decreased intention to adopt protective measures and increase their future privacy risks [28].

"As far as I know, there haven't been any court cases resulting from this kind of service [VATs]. I'm sure eventually there will be [...]. Until about five or six years ago, there weren't too many court cases involving Twitter and Facebook, but now the courts are full of them. The blind and low vision community is pretty small, but I'm sure at some point, there will be some kind of case where someone's information has been shared." In this section, we present the participants' *predictions*, *desires*, *tolerances* and what they believe to be their *rights* for visual privacy-preserving VATs according to (1) Their values related to privacy-preservation, and (2) Along the data life-cycle.

4.2.1 Value-Centered Visual Privacy Expectations According to Recipient Type. We present the value-based expectations according to two recipient types: (1) Remote sighted assistants, and (2) Software developers working or third parties involved in the creation and maintenance of AI-based VATs. The recipient type is a significant contextual factor influencing privacy [93, 94, 113], and recipients may be exposed to liability [63, 73]. The Supplementary Materials contain a table summarizing these findings.

Remote Sighted Assistants. **Control.** Participants expect: (1) to have access to all data they create and share with VATs, and (2) To delete that data from the VAT's servers. They had a low tolerance for restrictions that limit their control over their own data (e.g., *"If there is a video, it needs to be made immediately available to me, since I'm paying for the service. Right?... Is it Aira's record or my record? And that should be defined, too! If there is truly visual privacy, it [image/video data] should be mine. Period. And therefore it should be mine, and I should have free use of it"* [P13]. Payment for service increases participants' sense of *personal data ownership*, an important topic in terms of data regulation (e.g., [4, 40, 41, 121]).

Trust and Confidentiality. Payment for service also increases participants' trust in remote-sighted assistants' commitment to preserving visual privacy. Describing why they trusted Aira, [P14] noted, *"If there's some kind of problem you [paid remote sighted assistant] have, like you didn't adhere to the terms of your employment, well then, 'bye-bye and you'll be hearing from our lawyers'."* In contrast, participants expressed low expectations of trust for volunteer remote-sighted assistants. Participants like [P15] also predicted volunteers offered less confidentiality, requiring information management labor from VAT users: *"I do trust Be My Eyes, but it's a qualified trust in the sense that you have to do your due diligence on your end as well. If I receive a piece of mail that I think will have the number or my credit card number, I might not want to run that by them. Just have them look at the outside of the envelope."*

Accountability and Transparency. Participants specified they expect remote-sighted assistants to be accountable to VATs' privacy policies (particularly when payment is involved). For instance, [P03] said, *"If you talk to anybody on the team, everybody should be able to answer that [questions about how private visual content data is processed], or tell you where to find it."* Participants expect remote-sighted assistants to provide notice in comprehensible formats.

Software Developers Working for AI-Based VATs. **Control.** Participants expressed low tolerance for developers accessing their images/videos and expect to be provided with choices to limit developers' access (e.g.,*"I would never give a developer access to my images without my permission, I would sue the crap out of them because that's not okay. That's right, there is a violation of privacy"* [P06]. When software developers do access their data, participants described that *"You'd want the data anonymized somehow"* [P04] and robust data security measures to be in place.

Trust and Confidentiality. Whereas payment increased participants' expectations of trust and confidentiality for VATs that use remote-sighted assistants, non-disclosure agreements were key for trusting software developers and third parties with their visual data. For example, [P08] stated, *"Sign a contract, sign an agreement not to share!"*, while [P10] said, *"I would like them to sign a privacy agreement and then have a copy of that, to know they actually agreed to keep things private. Also, if a person does JAWS scripting and looks at sensitive information, he must sign a privacy agreement."* [P10] invoked their experience with other access technologies to define their expectations for VATs.

Accountability and transparency. Participants expect software developers to sign non-disclosure agreements, but also expect VAT companies to hold software developers accountable to these agreements and not use visual data outside contractual roles, and that sharing, leaking, or selling of users' images/videos is prohibited. For example, [P01] said,*"If you access people's private information, you're not supposed to share it. And if you do, you could get fired or whatever"* [P01]. Participants also expect the same for third parties, such as *"Have the privacy agreements included third parties or for the software developer to take responsibility for what their third party counterparts do"* [P12].

4.3 Expectations for Privacy-preserving VATs According to Stages of the Data Life-cycle

Here we present participants' expectations for how VATs should convey their practices of visual privacy preservation, and the nature of those privacy practices throughout the *data life-cycle*—the passage of data from initial generation to deletion. We highlight the privacy-preserving mechanisms and features participants expect from VATs, particularly when remote-sighted assistants and software developers access their private visual content. The Supplementary Materials include a table summarizing these findings.

4.3.1 Conveyance of VATs' Privacy Practices. Privacy policies and terms of use agreements are mechanisms companies use to convey to users their protocol for consumer data handling [106, 112]. Though participants had expectations of themselves to be informed of VAT companies' privacy policies (e.g., *"It's the end-user's responsibility to know what they're getting into and to read that privacy policy. Some of these privacy and legal agreements are pages and pages long, but you still need to try to read them"*, like most technology users [80, 81, 86, 87] they also have little tolerance for reading privacy policies (e.g., *"No way, I don't read them. Perhaps [I would] with a more understandable, tighter privacy policy written in plain simple language while drawing user's attention"* [P14]).

To more effectively communicate and get their attention, participants suggested using alternative formats (e.g., *"Maybe there could be ways to get more information about unfamiliar terms or key terms"* [P07]), such as in-app reference materials, to help understand the terminology used to describe the function of AI-based techniques. Participants' desires for the delivery of notice also diverge from traditional privacy policy formats to include more personalized

real-time approaches. For example, without specifying the type of VAT, [P10] described their desire for *"An email or pop-up in the app to show you like, okay, this is what happens to your images, and how long we keep them or where they go"*. Participants also wanted to be informed when VAT's privacy policies are updated (e.g., *"It is definitely important to know when any changes are made to the privacy policy"* [P15]) including changes in where their data is stored (e.g., *"Who knows whether the data center in the country where it's being kept is secure?"* [P14]), an important consideration as different countries follow different data regulations [29]. They also wanted to know if and how remote-sighted assistants were trained (e.g., *"My son signed up to be a volunteer with Be My Eyes, and he said all he had to do was read a little booklet. I would like to see that [what they learn about private content]"*[P08]) or if software developers followed confidentiality agreements.

4.3.2 Data Collection and Storage. Participants' expectations for privacy-preservation during data collection greatly centered on *access controls* for sessions with remote-sighted assistants, and for reduced–if not eliminated–collection of their data containing private visual information for AI-based VATs. Participants had clearly different expectations for remote visual assistants vs. developers.

Remote Visual Assistants. Participants wanted to limit remote sighted assistants' view. One idea involved a digital curtain that could quickly turn off the user's camera when somebody comes into the background, thus mitigating the risk of multiparty privacy conflicts (e.g., *"There needs to be a button to just make the screen go black. I mean, if the five-year-old walks in, you can just darken the screen and turn it back on again when they leave. Just like having a Screen Curtain [an iOS VoiceOver privacy feature]"* [P06]. A second idea was to enable users to zoom into specific content or crop irrelevant content from the sighted assistants' view (e.g., *"Maybe there is a part of the app where you can zoom in and out on your camera, that way you have more control of what your camera is actually seeing. And maybe pick a range. Have a little picker on it and say, I want this image, or I want my camera to only see three feet by five feet"[P12].*[16] As a backup, participants also shared the desire to immediately delete images/videos they had shared for visual assistance, so they do not have to submit a help desk ticket. One complexity identified with deleting these images/videos, however, was that remote visual assistants could save images with private visual content on their personal computers (e.g., *"If they do [use a laptop], are they securing that? Because inherently home PCs will be less secure [than a cloud server]"*[P04].

AI-Based VATs and their Developers. As described earlier, participants expressed little tolerance for AI-based VATs collecting images or videos that contain private content or providing developers access to that content, (e.g., *"Leave it on my device. Don't upload it to your system"* [P05], *"Just make sure it stays on my device"* [P08]). If data is collected, participants expect VATs to *"Dump it after use. Destroy it. Send it to data heaven"* [P06]. The participants predicted that VATs may use the data they collect to train their systems. In the words of [P11], *"I would really, really hope that that image is only accessed temporarily and immediately destroyed thereafter. I know there is stuff to be said about machine learning, trying to improve and review, but I'd rather they immediately destroyed."* At the same time, participants expressed a range of tolerances for their data being collected for the development of VATs. Some wanted no data collected, while others were open to visual data collection if private content is first removed, and if VATs are transparent about the uses of that data, and enacted strong security protocols (e.g., *"I would like to keep as many of their pictures on their device as possible. If there was a picture that wasn't sensitive, and I was okay with them using it for the database, maybe with an electronic consent form and a security measure"* [P06].

4.3.3 Data Processing. Processing visual data collected from individuals who are blind has proven application to the development of training remote sighted assistants [75, 110] and machine learning models that generate automated descriptions (e.g., [125]). Given the differences in how humans vs. AI interprets and describes visual information, participants shared different expectations for VAT employees offering visual assistance and those developing visual assistance.

Remote Sighted Assistants. Fundamentally, participants expect remote-sighted assistants to learn to recognize and notify users when they see private content in an image/video. For example, [P06] said, *"I would want them [the sighted assistant] to be honest about whether there's somebody inadvertently wandering into the frame, a person who doesn't need their picture shared on social media without their consent or knowledge. Just say 'Hey, just a heads-up, somebody happened to walk by when we took that picture. Maybe we want to take another one, or 'Are you aware there are people in the frame?' Just let me know, because then I will think about it, 'Oh, yeah. Maybe I shouldn't share that one'"*[17]. Participants also shared the desire for sighted assistants to be sensitive in their communication about such content, especially with stigmatizing categories such as guns, sex toys, or medications. For instance, [P07] shared, *"When you're letting someone know [that there is private content in their image/video], that could be a really sensitive conversation. Someone could feel embarrassed, or maybe that would come with reassurance. 'We're going to be encrypting and then getting rid of this content'"*.

AI-Based VATs and Developers. Whereas participants' expectations for remote sighted assistants focused on the delivery of information about the presence of private visual content, their expectations for privacy preservation from AI-based VATs focused on recognizing and obscuring the content from view (e.g.,*"Just blur out any really private information in the picture [...] like what Zoom does where you put the background above your head so no one can see you're crazy messy office"* [P05]. Similarly, when [P07] said *"Get rid of the private information"*, they independently predicted obfuscation as a possible privacy-preserving technique. Others made more lofty predictions focused on processing images so that private visual content would be omitted from the visual assistance they would receive (i.e., descriptions and visual question answering). *"They [VAT developers] could introduce some modifications to the video feed to focus on common objects, papers, and brands that users might need assistance with. And for that reason, AI could be used to alter the*

[16]The feature would enable users with some vision to remove portions of the live video feed. [P13] did not discuss how to make this feature accessible non-visually.

[17]This participant's reference to another person as the object of disclosure exemplifies a third person effect–an indicator of a bias or optimistic perception that can result in people's decreased intention to adopt protective measures and increase their privacy risks in the future [28].

video feed slightly [for that reason]" [P12]). To our understanding, visual access to image/video content would be reduced by omitting a description of that content, though the possibility of selecting the types of content a person wants to be described for machine training has been explored [60]. As blind people's images/videos have been used in the past for training algorithms (e.g., [53]), participants predicted that private visual content might be in the images/videos processed for training algorithms, and had low tolerance for the use of *real* private visual content to be used for training recognition and obfuscation-based models (e.g., *"I was having a discussion about something with photo quality or something [...] for evidence of their apartment being broken into [...] Could you create synthetic data that would actually create a good enough model? Cause God knows. It shouldn't be trained on people's actual private information"* [P07].

4.3.4 Data Security Management. In addition to the access controls identified above, they shared expectations for security protocols when their data is stored and processed *on the cloud* (i.e. through server-based frameworks) and *on-device* (i.e. frameworks deployed on portable platforms with a focus on smartphones).

Cloud-Based. For VATs that store or process images/videos *on the cloud*, participants expressed a range of data security management expectations related to data minimization, data encryption, and data deletion. Data minimization is the requirement for a system to only retain the user data necessary to deliver service [21] or in the words of [P16] *"...only data necessary to complete the task. Nothing that would be identifiable as far as to identify the person's name [...] There are situations where their location might be needed, but if you can get around that, that would be nice."* Encryption offers a well-known solution to making the content indecipherable in the case a user account is breached [32], though when images/videos are stored on the cloud, encryption is highly complex [68]; [P14] specifically discussed the complexity of encryption with volunteer-based VAT services that involve personal devices, stating *"As I understand it, the app [Be My Eyes] uses encryption to share data between your phone and the volunteer's phone, but it is not a peer-to-peer network, so it has to go through a server somewhere, as well on from my phone to the volunteer's phone. So, what happens to it there... I mean, are they using it appropriately? Are they protecting it appropriately? Is it even encrypted?"*

Device-Based. Participants had greater tolerance or fewer stipulations for device-based processing, and viewed this approach as more secure (e.g., *"I think an on-device option would be good. And if your storage doesn't cut it, then easy online [processing] and then deleted"* [P05]). As exemplified by [P05] and above, participants expect that if VATs use device-based processing, their data should not be collected, and if so, immediately deleted.

4.4 Reactions to AI-Based Recognition and Obfuscation

After learning about participants' baseline expectations for privacy preservation when using VATs, we presented possibilities for how AI-based techniques could better preserve visual privacy in the future. We asked participants to consider future VATs that could automatically identify private objects within images/videos and notify users of that risk, and (2) obfuscate that private content. While some participants had already independently introduced obfuscation earlier in the interviews, to ground all participants' reactions to these ideas, we asked them to imagine using the techniques in a given scenario, where they are: *cleaning their home and wanting to access a medical document they find*—information typically kept private in the United States through the Health Insurance Portability and Accountability Act of 1996 (HIPAA) [42].

4.4.1 Private Object Recognition. Current research identifies *object recognition* as a viable solution, though not yet accurate or necessarily fair, for identifying content in blind people's images (e.g., [19, 53, 83]); this AI-based approach has yet to be used for the recognition of private visual content in blind people's images. Here, we present the advantages and disadvantages of these techniques, as identified by VAT users. We also share the choices and features they would want if the technique were deployed.

Advantages of Private Object Recognition. Participants predicted that recognition of private objects would increase awareness of their surroundings if accurate (e.g., *"I am not always keeping track of what I am focusing on...like, if I'm looking at a certain part of the room and I completely turn around, then I would be concerned about what other objects I am exposing the camera's view too. And therefore, if the app could let me know (if I could count on the app to let me know)."*[P11]). Participants also predicted the technique could learn about their personal preferences for visual assistance by training it on examples of personal private content: (e.g., ,*"It could be an interactive learning program where I would say, 'Okay, this should be private'. As it describes the image to me, I would give it active feedback so it can learn what's private to me"[P14].*

Disadvantages. While recognizing the benefits of automation, participants also added several caveats, including the impact of the feature on their personal management practices, including their sense of responsibility to manage their own private information (e.g., *"It just sounds like so much effort [for the system] to try to anticipate every possible future contingency when what you really just need to do is clean up your act. I'm very old school that way, sorry"* [P016]), as well as the loss of valuable human connections experienced when interacting with remote sighted assistants: (*"I hate those automatic bot things [...] Oh, for God's sake, just let me talk to a person. So, I can see myself getting really annoyed by the bot telling me...I'd rather have a human being say, 'Heads up, you've got your credit card in the frame', or whatever it happens to be"*[P06].

In alignment with [14], participants also considered the potential impacts of algorithm errors, including both false positives (i.e., recognizing something as private when it is not) and false negatives (i.e., not recognizing something as private when it is). Participants also observed that visual privacy is contextual to the individual, so the minimum viable approach to create universal baseline descriptions (i.e., [114]) may not deliver the visual assistance they expect or need. Culture, religion, gender, prior visual experience are examples of specific factors that likely influence VAT users' preferences and use of obfuscation-based features [14].

Third, the participants focused on potential technical and hardware limitations that could affect the user experience. For VATs that process data *on their devices*, participants indicated that AI-based systems use a lot of battery power to operate the models on the device. In the words of [P14], *"The main con with on-device processing*

is the power of your device. The ability to run this type of algorithm." Another concern was how this additional private content check could impact latency. For example, as [P11] shared, *"It would require solid latency so every video frame is first given to the AI [on device] and determine whether there is any private information or not. Only if there isn't, then that feed is transmitted. And if there is, it releases a prompt. All of that seems to introduce too much complexity and slow down."*

Additional Design Considerations. The participants shared two additional expectations for privacy-preserving features to be included in VATs that use object recognition.

Content Description: Blind people want information about the content recognized for obfuscation, before and after the action [14]. Participants in our study shared expectations for the obfuscation-based systems to communicate the presence of private content through an earcon (making a sound), naming the type of content (e.g., medical information or medical bill), or both, or by providing details about the composition of the image or video keyframes. For example, [P08] specified they would want to know *"What the information is and maybe where in the image it is."*

Participants enumerated several options for how they would want to proceed once private visual content is flagged: (1) Stop the app from sending the image/video so the user can either retake it or abort, (2) Remove the private content from the image/video through 'cropping' or 'deleting'[18], or (3) Provide additional choices for where the images/videos will be sent based on how private the user deems the content. For example, [P05] said, *"You could just go 'it's private,' or 'this looks a bit private. Are you sure you want to do this? Are you really sure?' And then if you click proceed, so it'll tell you what it [the private content] is."*

Supports for Image/Video Capture and Sharing: Participants also suggested VATs provide photography support for the user to 'retake the photo if private content is shown' or push notifications that would enable them to specify 'no obfuscation; proceed with full description' or 'proceed with obfuscation'. Describing the recipients choices, and including recipients outside the VAT service, [P01] said, *"Maybe you've got check boxes, 'do you want this information kept only at your primary doctor's office, yes or no? Do you want this information to be put in your general medical records, yes or no?... And this might change if I start working with some sort of system, but at this point, I would want as much choice as I could get."* Participants also predicted situations when they would want to share private content, or turn off the recognition feature at will.

4.4.2 Obfuscation. We asked participants to consider the use of automated obfuscation, defined as *"A feature that uses AI to identify and automatically black-out or blur private visual content before you share the image with a sighted visual assistant."*

Advantages of Obfuscation. Alhablri et al. (2022) [15] identify three use cases in which blind people expect obfuscation to be useful: to gain control over text, manage impressions, and enhance the visual description. The participants in our study similarly expect enhancement of visual description but emphasized the use of the technique to remove private visual content from the background of their images/videos[19]. They predicted descriptions would be delivered more readily when private visual content is removed (i.e., enhancement of visual description [15]) as one use case for obfuscation (e.g., [P04] shared, *"If it was stuff in the background, it could cut that out, whether you'd like I said, I'd cropped to focus the thing. Just cut the background, because it's extraneous. That would be important for third-party protection"*. Here, the participant sees the technique as offering greater protection when the image/video is shared outside their control and the visual assistance technology.

Disadvantages. Despite these potential advantages, participants also discussed how accuracy and processing delays could interfere with their access to information. As in [14], they predicted that automated obfuscation could obscure the wrong content and thus create an obstacle to the visual assistance they need. As with object recognition, participants also identified that obfuscation could add (1) Inefficiency due to additional data processing, for example, saying *"It will take a little longer"* [P10], especially in the context of videos (e.g., *"It seems a bit resource intensive to keep track of that private object as the video feed moves around a little bit..."*[P12], (2) loss of agency and awareness (e.g., *"It could be a problem because then I don't know what's going on [what content it blacks out or not"* [P04], and (3) reduce their sense of responsibility and knowledge about how to take images/videos in privacy-preserving ways: [P14] shared, *"It's like, 'how could I edit the picture on my own, like if I was fully sighted? How close can that come to that? Because it should be similar. Like if I was sighted. And you could make those modifications, you know?"*

Additional Design Considerations. The participants also shared three additional expectations for privacy-preserving features to be included in VATs that obfuscation.

Black-Out or Blur: When hearing about the feature, participants asked for clarification about 'blacking out' vs. 'blurring' (e.g., *"I don't think I understand the difference"*[P08]), while others shared their rationale for using one approach over another. Most participants indicated they would prefer the content blacked out as part of a security precaution since *"Blurring can be undone*" [P06], which prior work has evidenced [78]). [P13] shared, *"There are people out there who are smart enough. They could un-blur it probably easier than they can un-black it out."* All the while, they expect to choose whether the obfuscation blacks out or blurs the private content, and to what extent (*"I would want to have control over whether it's still blacked-out or not, or if there's a way to black-out another part or something or only certain parts"*[P04]), and indicated their selections for the type of obfuscation would say, depends on the type of private content, such as *"I would say, depends. Like nudity should be blurred, so they [the visual assistant] could say, 'Hey, this has an image of a person, and the middle of it has been blurred out. It's probably like a nude photo or something, so at least they could give you a vague description of what it is"* [P01].

Controls: Fundamentally, participants shared their desires for a feature that would allow them to turn obfuscation on/off, and specify when to use it and for which private content types. For example, [P01] said *"...to control what it does and doesn't block, would be useful."* They also want the system to explain how it determined that prediction (e.g., *"Read to me what is identified [from the image]*

[18]Participants did not identify blacking-out or blurring private content as obfuscation possibilities before the research team introduced these ideas.

[19]Private visual content appears in the background of blind people's images [52]

and what it [the AI] is thinking is private and why and then say, 'We blocked out this...' [the VAT should] read exactly what it's blocked out" [P04]. Those who wanted additional details about the attributes and characteristics obfuscated private content, and shared statements such as *"Well, if you can't get as much detail as possible, then there's not much point in it."* [P10], akin to providing an overview first, recognizing and identifying the salient objects, and if necessary, providing additional details.

Evaluation: Participants noted that having an *accuracy threshold* would help them feel more confident about using obfuscation techniques (e.g., *"If you could go in and tell it 'Don't turn off text blocking, or whatever, but block offensive content. You could toggle that on and off too, depending on how accurate it is."* When asked directly whether they would feel comfortable using a system that produced accurate results less than 100% of the time, we obtained a variety of responses ranging from, *Yes* [I would be comfortable] (e.g., *' 'I don't think you can ever reach a hundred percent"*[P13]—specifying they would still want the feature if it was accurate 70% of the time), to *No* [I would not be comfortable] and would only use the feature if it was accurate 100% of the time (e.g., *"If it's saying it's not going to be very accurate, then I would be less likely to want to use it because I couldn't trust it."*[P04]). Across all participants, the average accuracy threshold they expect is 88.75% of the time (SD 9.92). Accuracy of computer vision algorithms is an important topic related to responsible AI, which has been explored in prior work related to blind people's expectations for image description (e.g., [19, 57]), but not with respect to the recognition or obfuscation of private visual content for this population.

Variations: At the end of the task, we asked participants whether the *information type* (e.g., medical information) presented within this scenario would change their perspectives on AI-based techniques, and all 16 participants shared sentiments such as *No* (e.g., *"No, I don't think so. I think I'd want that heads-up in the same situation, regardless of what the information is"*[P08]. All the while, they acknowledged that some private visual content types may be more sensitive to them than others, according to who receives the image/video and in what contexts they are used. In addition, when asked if their perspectives would change, whether the private visual content was in an *image* or a *video*. Only three participants said *Yes* (e.g., *"Because if there's no information in there when I start the video, it doesn't mean there's no information in there when I finish"*[P06], which means processing would need to occur for multiple frames in the video). A table in the Supplementary Materials provides a summary of the findings in this section.

5 DISCUSSION

The findings show that blind VAT users have stricter value-centered expectations for software developers and third parties working for VATs than remote-sighted assistants. Trust can be established with remote-sighted assistants during direct interactions (mediated by users' own privacy-preserving information management behaviors), while with software developers these interactions do not occur. Participants had cybersecurity and data-protection-oriented questions about the underlying technical infrastructure for VATs that collect their data, which would be cause for concern, as different countries follow different data regulations [29]. All the while, participants expect that payment for service ensures greater confidentiality, impression management, and ultimately trust with both types of services. They expect non-disclosure agreements anytime a software developer accesses their visual data. The Supplementary Materials provide a bulleted overview of the user-identified strategies VAT companies may use to align their service offerings with users' visual privacy expectations. Here provide considerations for further consideration.

5.1 Extending Privacy-Preserving Practices Offered and Supported by VATs

Users expect VATs to inform them about the handling of images/videos in a transparent and interactive manner, through channels other than privacy policies that simply do not work for most people [44, 72]. Regarding transparency, participants wanted notice about where their data are stored and when that location changes. Participants also expect to have easy access to records of what happens to their data. This is akin to *datasheets for datasets* [45] but would be available to individual users and for single images/videos. Delivery of this information should be as close as possible to the time of visual data collection, with options to permanently dismiss—*"Don't tell me again"*—such notices until an update to the privacy policy occurs.

Participants expect that payment for service increases visual-privacy preservation. VATs could offer a tiered approach, where users pay more or agree to share their data to support further system improvement while gaining more control over that data (e.g., choices and personal settings to manage and automate privacy preservation throughout the data life-cycle). However, such an approach could quickly lead to further inequities for VAT users in socially, economically, and educationally marginalized conditions, especially since blind people are systematically underemployed [95] and the cost is already an inhibiting factor for blind people wanting visual assistance [111]. Given that VAT users highly value control of their data, acting as one's own data broker could be an alternative to shift power and offset the requirement to share personal data for access to social services—what one participant referred to as a *"crisis of privacy"* [P16]. Importantly, with this recommendation, a core challenge remains: how to non-visually identify whether visual content that is safe to share.

5.2 Obfuscation

In the future, if VATs use image/video recognition and obfuscation techniques, such functionality needs to be humanizing, disability-first, and lightweight. Participants envisioned successful privacy-preserving VATs with obfuscation capabilities to enhance control, independence, and security. As prerequisites, they have low latency and be *trained on blind people's data*. Prior work observed that training object recognition models with images/videos captured non-visually by blind people outperform standard object recognition models for non-visual accessibility applications (e.g. [52, 53, 84]). Participants shared their prediction that obfuscation techniques will produce inaccurate results—through false negatives and positives, and obscuring visual assistance (also found in [15])–if not trained on disability-first datasets [118]. A disability-first approach advocates for the collection of data that serves a disability community first, and then could be generalized [118]. Sharma et al. [108]

recently collected the first dataset of images/videos containing private visual content captured by blind people. VATs that employ obfuscation as a privacy-preserving measure may draw on this and other disability-first datasets for AI-based obfuscation, to meet initial user expectations of inclusion and leadership—if not more accurate recognition and obfuscation.

More generally, participants reflected that heavy dependence on AI-based processing would reduce their sense of personal responsibility for information management. Alhabri et al. [15] observed that blind VAT users view obfuscation as a task that requires *interdependence* [18] in part due to the nuances that gender, religion, and ethnicity instill in a user's privacy decision-making. Like the participants in our study, Albarbi's participants expected to provide input at every stage of the obfuscation process: *"Without opportunities to interject through dismissing/consenting obfuscation decisions, blind people expect intrusive or unfair results"* [15]. More work is needed to understand blind people's tolerance for hand-offs [92] between humans and AI in the field of privacy-preservation research and engineering. Best next steps will involve blind users in the design and evaluation of a working prototype to observe their decision-making about obfuscation, including what information resources they need to understand visual concepts like blurring and inpainting, if and when obfuscation had been initiated, and how to invoke feelings of trust when obfuscations are applied.

5.3 User-Led and Personalized Privacy Preservation

Recognizing the importance of interdependence motivates exploration of how humans can off-load privacy preservation to AI in ways that respect their values and expectations. Participants in this study predicted the challenge of delivering personalized privacy-preserving features, including overcoming the loss of personal responsibility and control that some users may experience, and the inability of AI-based techniques to accurately account for the contextual nature of privacy [93]. They expected automated systems to have difficulty determining what a user considers private for a context (e.g. task, location, gender, religion, ethnicity, and others); privacy-focused decision-making is notoriously inconsistent [3].

At a time when research on automation of visual assistance has just begun to explore contextually-aware visual assistance [71, 114] and advanced AI-in-the-loop approaches (e.g., [65, 126]), additional research is needed to address algorithmic inaccuracies that participants expect from object recognition within images (e.g., [52]), examine if taxonomies of private visual content (e.g., [77, 112]) can generalize, whether user-specified metadata about scenarios in which they consider image/video content to be private is useful for personalization, and how images/videos with obfuscated content can be described to provide privacy-aware decision making while providing complete visual access.

As a step towards establishing roles in the process of handing off labor to AI, efforts have been made to include blind people as curators of image/video datasets for training object recognizers (e.g., [83]) and developing accessible interfaces that enable blind people to train teachable object recognizers [60, 64]). These efforts have not been used to flag private visual content or provide subsequent steps needed for privacy preservation. Moving forward, more research is needed to explore how the aforementioned approaches may apply to visual privacy preservation. Efforts that build on prior research related to remote visual assistance (e.g., [75]) may explore the direct interactions users have with remote-sighted assistants, the visual questions they ask, the choices desired, and more to identify other tasks to hand off to AI for privacy preservation vs. the tasks that can only be addressed by human-sighted assistants.

5.4 Limitations

We had several reflections on the study design during implementation. Though saturation was reached regarding the expectations shared in this paper, our attempt to appropriately scope the study likely diminished the diversity of perspectives represented in the final dataset. The sample is limited to participants from the USA and England who currently use VATs, who were recruited from a list-serve of people who make themselves available to "test and evaluate the accessibility of websites, devices, appliances, equipment, and other products and services" [95] (and thus have a high baseline of digital, including visual and privacy literacy). In turn, we inadvertently excluded participants from (1) other countries, (2) former VAT users who have given up on VATs, or (3) Potential VATs users who declined due to concerns about the technologies. Our protocol did not include questions about participants' *visual literacy* or *privacy literacy*, unsure whether this factor affected participants' expectations. That said, we retrospectively reflect that participants demonstrated declarative knowledge and procedural knowledge—two important indicators of privacy literacy [99] [20], and understanding of the practice. We did not collect behavioral data, and thus our approach may not fully reflect blind people's privacy expectations [2].

6 CONCLUSION

We contribute novel findings about ($N = 16$) totally blind participants' visual privacy expectations when sharing personal and potentially private information with visual assistance technologies (e.g., Aira, Be My Eyes, Seeing AI, and Envision) that receive users' images/videos/live-feed to provide visual assistance. In a departure from prior work that has focused on studying privacy *concerns*, we instead focus on *expectations* as a generative approach to identifying the desires, predictions, tolerances, and ideas individuals who are blind have about what they deserve with respect to visual privacy from visual assistance technologies. Three overarching expectations: (1) never collect images/videos that contain private visual content; (2) develop interactive features that provide logs, in real-time, of who handles their visual data and why, while providing non-disclosure agreements from software developers and third parties; (3) practice caution and ethical consideration during the development and deployment of privacy-preserving AI. We discuss the challenges of developing AI-based privacy-preserving VATs based on these expectations and other ethical considerations including how to train privacy-preserving machine learning models that are inclusive of blind people's data, without revealing their private visual content.

[20]Declarative understanding of regulations and laws of online private visual data protection, and the skills to convey a breadth and depth of knowledge on what happens to users data as soon as data is created until it has been sent to data heaven. Procedural knowledge pertains to the use of privacy-preserving protective technologies cite [123]

ACKNOWLEDGMENTS

We thank our participants for their contributions and for sharing their insights. We also thank Chancey Fleet for her initial consultation on the topic of this research. This research was made possible by the NSF/CRA CIFellows program, and NSF Grant #2148080.

REFERENCES

[1] 2021. Envision App. https://www.letsenvision.com/envision-app
[2] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, and others. 2017. Nudges for privacy and security: Understanding and assisting users choices online. *ACM Computing Surveys (CSUR)* 50, 3 (2017), 1–41. Publisher: ACM New York, NY, USA.
[3] Alessandro Acquisti, Laura Brandimarte, and Jeff Hancock. 2022. How privacy's past may shape its future. *Science* 375, 6578 (Jan. 2022), 270–272. https://doi.org/10.1126/science.abj0826 Publisher: American Association for the Advancement of Science.
[4] Andria Agesilaou and Eleni A. Kyza. 2022. Whose data are they? Elementary school students' conceptualization of data ownership and privacy of personal digital data. *International Journal of Child-Computer Interaction* 33 (Sept. 2022), 100462. https://doi.org/10.1016/j.ijcci.2022.100462
[5] Maleeha Rafiq Ahmed and Muhammad Asif Naveed. 2019. Seeing Beyond Sight: The Academic Information Behavior of Visually Impaired Students. *Pakistan Library & Information Science Journal* 50, 2 (2019).
[6] 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 - CHI '15*. ACM Press, Seoul, Republic of Korea, 3523–3532. https://doi.org/10.1145/2702123.2702334
[7] T. Ahmed, R. Hoyle, P. Shaffer, K. Connelly, D. Crandall, and A. Kapadia. 2017. Understanding Physical Safety, Security, and Privacy Concerns of People with Visual Impairments. *IEEE Internet Computing* (2017), 1–1. https://doi.org/10.1109/MIC.2017.265103316 Conference Name: IEEE Internet Computing.
[8] 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. https://www.usenix.org/conference/soups2016/technical-sessions/presentation/ahmed
[9] Aira. 2020. Aira. https://aira.io/ Library Catalog: aira.io.
[10] Amaia Aizpurua, Myriam Arrue, and Markel Vigo. 2013. Uncovering the role of expectations on perceived web accessibility. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10.1145/2513383.2513411
[11] Taslima Akter. 2020. Privacy Considerations of the Visually Impaired with Camera Based Assistive Tools. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing (CSCW '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 69–74. https://doi.org/10.1145/3406865.3418382
[12] 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* (Feb. 2022). https://doi.org/10.1145/3506857
[13] 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)*. https://www.usenix.org/conference/usenixsecurity20/presentation/akter
[14] Rahaf Alharbi. 2018. Understanding Emerging Obfuscation Technologies in Visual Description Services for Blind and Low Vision People. (2018), 28.
[15] Rahaf Alharbi, Robin N. Brewer, and Sarita Schoenebeck. 2022. Understanding Emerging Obfuscation Technologies in Visual Description Services for Blind and Low Vision People. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 469:1–469:33. https://doi.org/10.1145/3555570
[16] Irwin Altman. 1977. Privacy Regulation: Culturally Universal or Culturally Specific? *Journal of Social Issues* 33, 3 (July 1977), 66–84. https://doi.org/10.1111/j.1540-4560.1977.tb01883.x
[17] Lemi Baruh, Ekin Secinti, and Zeynep Cemalcilar. 2017. Online Privacy Concerns and Privacy Management: A Meta-Analytical Review. *Journal of Communication* 67, 1 (Feb. 2017), 26–53. https://doi.org/10.1111/jcom.12276
[18] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Galway Ireland, 161–173. https://doi.org/10.1145/3234695.3236348
[19] Cynthia L. Bennett, Cole Gleason, Morgan Klaus Scheuerman, Jeffrey P. Bigham, Anhong Guo, and Alexandra To. 2021. "It's Complicated": Negotiating Accessibility and (Mis)Representation in Image Descriptions of Race, Gender, and Disability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3411764.3445498
[20] BeSpecular. 2020. BeSpecular. https://www.bespecular.com/
[21] Asia J. Biega, Peter Potash, Hal Daumé, Fernando Diaz, and Michèle Finck. 2020. Operationalizing the Legal Principle of Data Minimization for Personalization. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, Virtual Event China, 399–408. https://doi.org/10.1145/3397271.3401034
[22] 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. https://doi.org/10.1145/1805986.1806020
[23] Aditya Bodi, Pooyan Fazli, Shasta Ihorn, Yue-Ting Siu, Andrew T Scott, Lothar Narins, Yash Kant, Abhishek Das, and Ilmi Yoon. 2021. Automated Video Description for Blind and Low Vision Users. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3411763.3451810
[24] 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 - CHI '13*. ACM Press, Paris, France, 2117. https://doi.org/10.1145/2470654.2481291
[25] Stacy M. Branham, Ali Abdolrahmani, William Easley, Morgan Scheuerman, Erick Ronquillo, and Amy Hurst. 2017. "Is Someone There? Do They Have a Gun": How Visual Information about Others Can Improve Personal Safety Management for Blind Individuals. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Baltimore Maryland USA, 260–269. https://doi.org/10.1145/3132525.3132534
[26] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa Publisher: Taylor & Francis.
[27] Serap Buyurgan. 2009. The Expectations of the Visually Impaired University Students from Museums. *Educational Sciences: Theory and Practice* 9, 3 (2009), 1191–1204. https://eric.ed.gov/?id=EJ858923 Publisher: Educational Consultancy, Ltd (EDAM).
[28] Hongliang Chen and David Atkin. 2021. Understanding third-person perception about Internet privacy risks. *New Media & Society* 23, 3 (March 2021), 419–437. https://doi.org/10.1177/1461444820902103 Publisher: SAGE Publications.
[29] Rong Chen. 2021. *Mapping Data Governance Legal Frameworks around the World: Findings from the Global Data Regulation Diagnostic*. The World Bank. https://doi.org/10.1596/1813-9450-9615
[30] Tai-Yin Chiu, Yinan Zhao, and Danna Gurari. 2020. Assessing Image Quality Issues for Real-World Problems. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 3643–3653. https://doi.org/10.1109/CVPR42600.2020.00370 ISSN: 2575-7075.
[31] Markus Christen, Bert Gordijn, and Michele Loi (Eds.). 2020. *The Ethics of Cybersecurity*. The International Library of Ethics, Law and Technology, Vol. 21. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-29053-5
[32] Deepika Deepika, Rajnesh Malik, Saurabh Kumar, Rishabh Gupta, and Ashutosh Kumar Singh. 2020. A Review on Data Privacy using Attribute-Based Encryption. *SSRN Electronic Journal* (2020). https://doi.org/10.2139/ssrn.3606261
[33] Jasmine DeHart, Chenguang Xu, Lisa Egede, and Christan Grant. 2021. Proposing an Interactive Audit Pipeline for Visual Privacy Research. *arXiv:2111.03984 [cs]* (Nov. 2021). http://arxiv.org/abs/2111.03984 arXiv: 2111.03984.
[34] Descript. [n. d.]. Descript is an all-in-one audio/video editor and screen recorder that works like a doc. https://web.descript.com/
[35] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. 2020. Ask the Experts: What Should Be on an IoT Privacy and Security Label?. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 447–464. https://doi.org/10.1109/SP40000.2020.00043
[36] Pardis Emami-Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. [n. d.]. Privacy Expectations and Preferences in an IoT World. Thirteenth Symposium on Usable Privacy and Security ([n. d.]), 15.
[37] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. 2019. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–12. https://doi.org/10.1145/3290605.3300764

[38] eSight. [n. d.]. Electronic eyewear for the visually impaired. https://esighteyewear.com/

[39] Be My Eyes. 2020. Be My Eyes - See the world together. https://www.bemyeyes.com/ Library Catalog: www.bemyeyes.com.

[40] Martin Fadler and Christine Legner. 2022. Data ownership revisited: clarifying data accountabilities in times of big data and analytics. *Journal of Business Analytics* 5, 1 (Jan. 2022), 123–139. https://doi.org/10.1080/2573234X.2021.1945961 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/2573234X.2021.1945961.

[41] Roberta Fischli. 2022. Data-owning democracy: Citizen empowerment through data ownership. *European Journal of Political Theory* (July 2022), 14748851221110316. https://doi.org/10.1177/14748851221110316 Publisher: SAGE Publications.

[42] Center for Disease Control and Prevention. 2019. Health Insurance Portability and Accountability Act of 1996 (HIPAA). https://www.cdc.gov/phlp/publications/topic/hipaa.html hippa_1996.

[43] Batya Friedman, Peter H Kahn, and Alan Borning. 2008. Value sensitive design and information systems. *The handbook of information and computer ethics* (2008), 69–101. https://doi.org/10.1002/9780470281819.ch4 Publisher: Wiley Online Library.

[44] Alisa Frik, Juliann Kim, Joshua Rafael Sanchez, and Joanne Ma. 2022. Users' Expectations About and Use of Smartphone Privacy and Security Settings. (2022), 24. https://doi.org/10.1145/3491102.3517504

[45] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for Datasets. https://doi.org/10.48550/arXiv.1803.09010 arXiv:1803.09010 [cs].

[46] Cole Gleason, Patrick Carrington, Lydia B. Chilton, Benjamin M. Gorman, Hernisa Kacorri, Andrés Monroy-Hernández, Meredith Ringel Morris, Garreth W. Tigwell, and Shaomei Wu. 2019. Addressing the Accessibility of Social Media. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 474–479. https://doi.org/10.1145/3311957.3359439

[47] Cole Gleason, Amy Pavel, Himalini Gururaj, Kris M Kitani, and Jeffrey P Bigham. 2020. Making GIFs Accessible. (2020), 10. https://doi.org/10.1145/3373625.3417027

[48] Cole Gleason, Amy Pavel, Xingyu Liu, Patrick Carrington, Lydia B. Chilton, and Jeffrey P. Bigham. 2019. Making Memes Accessible. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Pittsburgh PA USA, 367–376. https://doi.org/10.1145/3308561.3353792

[49] Cole Gleason, Amy Pavel, Emma McCamey, Christina Low, Patrick Carrington, Kris M. Kitani, and Jeffrey P. Bigham. 2020. Twitter A11y: A Browser Extension to Make Twitter Images Accessible. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3313831.3376728

[50] 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*. ACM, 518.

[51] Danna Gurari and Kristen Grauman. 2017. CrowdVerge: Predicting if people will agree on the answer to a visual question. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3511–3522. https://doi.org/10.1145/3025453.3025781

[52] Danna Gurari, Qing Li, Chi Lin, Yinan 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 Conference on Computer Vision and Pattern Recognition*. 939–948.

[53] 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. https://doi.org/10.48550/arXiv.1802.08218 arXiv:1802.08218 [cs].

[54] Danna Gurari, Yinan Zhao, Suyog Dutt Jain, Margrit Betke, and Kristen Grauman. 2019. Predicting How to Distribute Work Between Algorithms and Humans to Segment an Image Batch. *arXiv:1905.00060 [cs]* (April 2019). http://arxiv.org/abs/1905.00060 arXiv: 1905.00060.

[55] Danna Gurari, Yinan Zhao, Meng Zhang, and Nilavra Bhattacharya. 2020. Captioning Images Taken by People Who Are Blind. In *Computer Vision – ECCV 2020 (Lecture Notes in Computer Science)*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer International Publishing, Cham, 417–434. https://doi.org/10.1007/978-3-030-58520-4_25

[56] Foad Hamidi, Kellie Poneres, Aaron Massey, and Amy Hurst. 2020. Using a participatory activities toolkit to elicit privacy expectations of adaptive assistive technologies. In *Proceedings of the 17th International Web for All Conference*. ACM, Taipei Taiwan, 1–12. https://doi.org/10.1145/3371300.3383336

[57] Margot Hanley, Apoorv Khandelwal, Hadar Averbuch-Elor, Noah Snavely, and Helen Nissenbaum. 2020. An Ethical Highlighter for People-Centric Dataset Creation. *arXiv:2011.13583 [cs]* (Nov. 2020). http://arxiv.org/abs/2011.13583 arXiv: 2011.13583.

[58] Hamza Harkous, Kassem Fawaz, Rémi Lebret, Florian Schaub, Kang G. Shin, and Karl Aberer. 2018. Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*. 531–548. https://www.usenix.org/conference/usenixsecurity18/presentation/harkous

[59] 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*.

[60] Jonggi Hong, Jaina Gandhi, Ernest Essuah Mensah, Ebrima H. Jarjue, Kyungjun Lee, and Hernisa Kacorri. 2022. Blind Users Accessing Their Training Images in Teachable Object Recognizers. https://doi.org/10.48550/arXiv.2208.07968 arXiv:2208.07968 [cs].

[61] Hsinfu Huang. 2018. Blind users' expectations of touch interfaces: factors affecting interface accessibility of touchscreen-based smartphones for people with moderate visual impairment. *Universal Access in the Information Society* 17, 2 (June 2018), 291–304. https://doi.org/10.1007/s10209-017-0550-z

[62] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. ACM, Denver Colorado USA, 781–792. https://doi.org/10.1145/2810103.2813603

[63] Leonardo Horn Iwaya, Gabriel Horn Iwaya, Simone Fischer-Hübner, and Andrea Valéria Steil. 2022. Organisational Privacy Culture and Climate: A Scoping Review. *IEEE Access* 10 (2022), 73907–73930. https://doi.org/10.1109/ACCESS.2022.3190373 Conference Name: IEEE Access.

[64] Hernisa Kacorri. 2017. Teachable machines for accessibility. *ACM SIGACCESS Accessibility and Computing* 119 (Nov. 2017), 10–18. https://doi.org/10.1145/3167902.3167904

[65] Rie Kamikubo, Utkarsh Dwivedi, and Hernisa Kacorri. 2021. Sharing Practices for Datasets Related to Accessibility and Aging. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3441852.3471208

[66] Rie Kamikubo, Naoya Kato, Keita Higuchi, Ryo Yonetani, and Yoichi Sato. 2020. Support Strategies for Remote Guides in Assisting People with Visual Impairments for Effective Indoor Navigation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376823

[67] Deborah Kent. 2002. Beyond Expectations: Being Blind and Becoming a Mother. *Sexuality and Disability* 20, 1 (March 2002), 81–88. https://doi.org/10.1023/A:1015238505439

[68] John S. Koh, Jason Nieh, and Steven M. Bellovin. 2021. Encrypted cloud photo storage using Google photos. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, Virtual Event Wisconsin, 136–149. https://doi.org/10.1145/3458864.3468220

[69] Stephen Kosack and Archon Fung. 2014. Does Transparency Improve Governance? *Annual Review of Political Science* 17, 1 (2014), 65–87. https://doi.org/10.1146/annurev-polisci-032210-144356 _eprint: https://doi.org/10.1146/annurev-polisci-032210-144356.

[70] Eliza Kostyra, Sylwia Żakowska Biemans, Katarzyna Śniegocka, and Anna Piotrowska. 2017. Food shopping, sensory determinants of food choice and meal preparation by visually impaired people. Obstacles and expectations in daily food experiences. *Appetite* 113 (June 2017), 14–22. https://doi.org/10.1016/j.appet.2017.02.008

[71] Elisa Kreiss, Cynthia Bennett, Shayan Hooshmand, Eric Zelikman, Meredith Ringel Morris, and Christopher Potts. 2022. Context Matters for Image Descriptions for Accessibility: Challenges for Referenceless Evaluation Metrics. https://doi.org/10.48550/arXiv.2205.10646 arXiv:2205.10646 [cs].

[72] Barbara Krumay and Jennifer Klar. 2020. Readability of Privacy Policies. In *Data and Applications Security and Privacy XXXIV (Lecture Notes in Computer Science)*, Anoop Singhal and Jaideep Vaidya (Eds.). Springer International Publishing, Cham, 388–399. https://doi.org/10.1007/978-3-030-49669-2_22

[73] Patrick Kühtreiber, Viktoriya Pak, and Delphine Reinhardt. 2022. A survey on solutions to support developers in privacy-preserving IoT development. *Pervasive and Mobile Computing* 85 (Sept. 2022), 101656. https://doi.org/10.1016/j.pmcj.2022.101656

[74] Kyungjun Lee, Daisuke Sato, Saki Asakawa, Chieko Asakawa, and Hernisa Kacorri. 2021. Accessing Passersby Proxemic Signals through a Head-Worn Camera: Opportunities and Limitations for the Blind. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event USA, 1–15. https://doi.org/10.1145/3441852.3471232

[75] Sooyeon Lee, Madison Reddie, Chun-Hua Tsai, Jordan Beck, Mary Beth Rosson, and John M. Carroll. 2020. The Emerging Professional Practice of Remote Sighted Assistance for People with Visual Impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu

HI USA, 1–12. https://doi.org/10.1145/3313831.3376591
[76] Sooyeon Lee, Rui Yu, Jingyi Xie, Syed Masum Billah, and John M. Carroll. 2022. Opportunities for Human-AI Collaboration in Remote Sighted Assistance. In *27th International Conference on Intelligent User Interfaces*. ACM, Helsinki Finland, 63–78. https://doi.org/10.1145/3490099.3511113
[77] Yifang Li, Nishant Vishwamitra, Hongxin Hu, and Kelly Caine. 2020. Towards A Taxonomy of Content Sensitivity and Sharing Preferences for Photos. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–14. https://doi.org/10.1145/3313831.3376498
[78] Yifang Li, Nishant Vishwamitra, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Blur vs. Block: Investigating the Effectiveness of Privacy-Enhancing Obfuscation for Images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, Honolulu, HI, USA, 1343–1351. https://doi.org/10.1109/CVPRW.2017.176
[79] Jialiu Lin, Shahriyar Amini, Jason I. Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, New York, NY, USA, 501–510. https://doi.org/10.1145/2370216.2370290
[80] Heather Richter Lipford, Jason Watson, Michael Whitney, Katherine Froiland, and Robert W. Reeder. 2010. Visual vs. compact: a comparison of privacy policy interfaces. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, Atlanta, Georgia, USA, 1111. https://doi.org/10.1145/1753326.1753492
[81] Fei Liu, Rohan Ramanath, Norman Sadeh, and Noah A. Smith. 2014. A Step Towards Usable Privacy Policy: Automatic Alignment of Privacy Statements. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Dublin City University and Association for Computational Linguistics, Dublin, Ireland, 884–894. https://www.aclweb.org/anthology/C14-1084
[82] Franco Loi, Arunan Sivanathan, Hassan Habibi Gharakheili, Adam Radford, and Vijay Sivaraman. 2017. Systematically Evaluating Security and Privacy for Consumer IoT Devices. In *Proceedings of the 2017 Workshop on Internet of Things Security and Privacy (IoTS&P '17)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3139937.3139938
[83] Sergio Mascetti, Mattia Ducci, Niccolò Cantù, Paolo Pecis, and Dragan Ahmetovic. 2021. Developing Accessible Mobile Applications with Cross-Platform Development Frameworks. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event USA, 1–5. https://doi.org/10.1145/3441852.3476469
[84] Daniela Massiceti, Luisa Zintgraf, John Bronskill, Lida Theodorou, Matthew Tobias Harris, Edward Cutrell, Cecily Morrison, Katja Hofmann, and Simone Stumpf. 2021. ORBIT: A Real-World Few-Shot Dataset for Teachable Object Recognition. 10818–10828. https://openaccess.thecvf.com/content/ICCV2021/html/Massiceti_ORBIT_A_Real-World_Few-Shot_Dataset_for_Teachable_Object_Recognition_ICCV_2021_paper.html
[85] MAXQDA. 2020. *All-In-One Tool for Qualitative Data Analysis & Mixed Methods*. https://www.maxqda.com/
[86] Aleecia M. McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society* 4 (2008), 543. Publisher: HeinOnline.
[87] Yannic Meier, Johanna Schäwel, Elias Kyewski, and Nicole C. Krämer. 2020. Applying Protection Motivation Theory to Predict Facebook Users’ Withdrawal and Disclosure Intentions. In *International Conference on Social Media and Society (SMSociety'20)*. Association for Computing Machinery, New York, NY, USA, 21–29. https://doi.org/10.1145/3400806.3400810
[88] Microsoft. 2020. Seeing AI App from Microsoft. https://www.microsoft.com/en-us/ai/seeing-ai Library Catalog: www.microsoft.com.
[89] Miles, Matthew. 2022. Qualitative Data Analysis. https://us.sagepub.com/en-us/nam/qualitative-data-analysis/book246128
[90] Valerie S Morash, Yue-Ting Siu, Joshua A Miele, Lucia Hasty, and Steven Landau. 2015. Guiding novice web workers in making image descriptions using templates. *ACM Transactions on Accessible Computing (TACCESS)* 7, 4 (2015), 12. Publisher: ACM.
[91] Meredith Ringel Morris, Jazette Johnson, Cynthia L. Bennett, and Edward Cutrell. 2018. Rich Representations of Visual Content for Screen Reader Users. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–11. https://doi.org/10.1145/3173574.3173633
[92] Deirdre K. Mulligan and Helen Nissenbaum. 2020. The Concept of Handoff as a Model for Ethical Analysis and Design. https://papers.ssrn.com/abstract=3784839
[93] Helen Nissenbaum. 2004. Privacy as Contextual Integrity. *Washington Law Review* 79 (2004), 41.
[94] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press. https://doi.org/10.1515/9780804772891
[95] National Federation of the Blind. [n. d.]. Blind Users Innovating and Leading Design | National Federation of the Blind. https://nfb.org/programs-services/center-excellence-nonvisual-access/blind-users-innovating-and-leading-design
[96] OrCam. 2021. OrCam MyEye, Now With the Groundbreaking Smart Reading Feature - YouTube. https://www.youtube.com/watch?v=bbEEmc0xtvw
[97] José Ramón Padilla-López, Alexandros Andre Chaaraoui, Feng Gu, and Francisco Flórez-Revuelta. 2015. Visual Privacy by Context: Proposal and Evaluation of a Level-Based Visualisation Scheme. *Sensors* 15, 6 (June 2015), 12959–12982. https://doi.org/10.3390/s150612959 Number: 6 Publisher: Multidisciplinary Digital Publishing Institute.
[98] Helen Petrie, Chandra Harrison, and Sundeep Dev. [n. d.]. Describing images on the Web: a survey of current practice and prospects for the future. ([n. d.]), 10.
[99] Christine Prince, Nessrine Omrani, Adnane Maalaoui, Marina Dabic, and Sascha Kraus. 2021. Are We Living in Surveillance Societies and Is Privacy an Illusion? An Empirical Study on Privacy Literacy and Privacy Concerns. *IEEE Transactions on Engineering Management* (2021), 1–18. https://doi.org/10.1109/TEM.2021.3092702 Conference Name: IEEE Transactions on Engineering Management.
[100] Ashwini Rao and Juergen Pfeffer. 2020. Types of Privacy Expectations. *Frontiers in Big Data* 3 (2020). https://doi.org/10.3389/fdata.2020.00007
[101] KNFB Reader. 2020. App features the best OCR. Turns print into speech or Braille instantly. iOS 3 now available. | KNFB Reader. https://knfbreader.com/
[102] Looktel Money Reader. 2020. *LookTel Money Reader for iPhone, iPod Touch and Mac*. http://www.looktel.com/moneyreader
[103] Joel R. Reidenberg, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T. Graves, Fei Liu, Aleecia McDonald, Thomas B. Norton, and Rohan Ramanath. 2015. Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding. *Berkeley Technology Law Journal* 30, 1 (2015), 1–88. https://heinonline.org/HOL/P?h=hein.journals/berktech30&i=51
[104] Ethan Z. Rong, Mo Morgana Zhou, Zhicong Lu, and Mingming Fan. 2022. "It Feels Like Being Locked in A Cage": Understanding Blind or Low Vision Streamers' Perceptions of Content Curation Algorithms. http://arxiv.org/abs/2204.11247 arXiv:2204.11247 [cs].
[105] Catherine E. Ross and Jaya Sastry. 1999. The Sense of Personal Control. In *Handbook of the Sociology of Mental Health*, Carol S. Aneshensel and Jo C. Phelan (Eds.). Springer US, Boston, MA, 369–394. https://doi.org/10.1007/0-387-36223-1_18
[106] Emma Sadjo, Leah Findlater, and Abigale Stangl. 2021. Landscape Analysis of Commercial Visual Assistance Technologies. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event USA, 1–4. https://doi.org/10.1145/3441852.3476521
[107] Evan Selinger, Jules Polonetsky, and Omer Tene. 2018. *The Cambridge Handbook of Consumer Privacy*. Cambridge University Press. Google-Books-ID: vnVSDwAAQBAJ.
[108] 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. (2023).
[109] Jiayu Shu, Rui Zheng, and Pan Hui. 2018. Cardea: context-aware visual privacy protection for photo taking and sharing. In *Proceedings of the 9th ACM Multimedia Systems Conference*. ACM, Amsterdam Netherlands, 304–315. https://doi.org/10.1145/3204949.3204973
[110] Rachel N. Simons, Danna Gurari, and Kenneth R. Fleischmann. 2020. "I Hope This Is Helpful": Understanding Crowdworkers' Challenges and Motivations for an Image Description Task. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 105:1–105:26. https://doi.org/10.1145/3415176
[111] 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 (CHI '20)*. Association for Computing Machinery, Honolulu, HI, USA, 1–13. https://doi.org/10.1145/3313831.3376404
[112] Abigale Stangl, Kristina Shiroma, Nathan Davis, Bo Xie, Kenneth R. Fleischmann, Leah Findlater, and Danna Gurari. 2022. Privacy Concerns for Visual Assistance Technologies. *ACM Transactions on Accessible Computing* 15, 2 (June 2022), 1–43. https://doi.org/10.1145/3517384
[113] Abigale Stangl, Kristina Shiroma, Bo Xie, Kenneth R. Fleischmann, and Danna Gurari. 2020. Visual Content Considered Private by People Who are Blind. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event Greece, 1–12. https://doi.org/10.1145/3373625.3417014
[114] Abigale Stangl, Nitin Verma, Kenneth R. Fleischmann, Meredith Ringel Morris, and Danna Gurari. 2021. Going Beyond One-Size-Fits-All Image Descriptions to Satisfy the Information Wants of People Who are Blind or Have Low Vision. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event USA, 1–15. https://doi.org/10.1145/3441852.3471233
[115] Supersense. 2021. Supersense - AI for Blind / Scan text, money and objects. https://www.supersense.app/

[116] Madiha Tabassum, Abdulmajeed Alqhatani, Marran Aldossari, and Heather Richter Lipford. 2018. Increasing User Attention with a Comic-based Policy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3173574.3173774

[117] TapTapSee. 2020. *Blind and Visually Impaired Assistive Technology - powered by CloudSight.ai Image Recognition API*. https://taptapseeapp.com/

[118] 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 *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, Virtual Event USA, 1–12. https://doi.org/10.1145/3441852.3471225

[119] Torrey Trust and Brian Horrocks. 2019. Six Key Elements Identified in an Active and Thriving Blended Community of Practice. *TechTrends* 63, 2 (March 2019), 108–115. https://doi.org/10.1007/s11528-018-0265-x

[120] Nguyen Anh Tu, Thien Huynh-The, Kok-Seng Wong, M. Fatih Demirci, and Young-Koo Lee. 2021. Toward efficient and intelligent video analytics with visual privacy protection for large-scale surveillance. *The Journal of Supercomputing* (May 2021). https://doi.org/10.1007/s11227-021-03865-7

[121] Takis Vidalis. 2022. Self-Ownership. In *The Emergence of Biolaw: The European Experience and the Evolutionary Approach*, Takis Vidalis (Ed.). Springer International Publishing, Cham, 173–180. https://doi.org/10.1007/978-3-031-02359-0_9

[122] Angelina Wang, Solon Barocas, Kristen Laird, and Hanna Wallach. 2022. Measuring Representational Harms in Image Captioning. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 324–335. https://doi.org/10.1145/3531146.3533099

[123] Christina L Wissinger and Penn State. [n. d.]. Privacy Literacy: From Theory to Practice. ([n. d.]), 12. https://doi.org/10.15760/comminfolit.2017.11.2.9

[124] 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.

[125] 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 (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 1180–1192. https://doi.org/10.1145/2998181.2998364

[126] 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* 29, 4 (Aug. 2022), 1–40. https://doi.org/10.1145/3501298

[127] Razieh Nokhbeh Zaeem, Rachel L. German, and K. Suzanne Barber. 2018. PrivacyCheck: Automatic Summarization of Privacy Policies Using Data Mining. *ACM Transactions on Internet Technology* 18, 4 (Nov. 2018), 1–18. https://doi.org/10.1145/3127519

[128] Sebastian Zimmeck, Ziqi Wang, Lieyong Zou, Roger Iyengar, Bin Liu, Florian Schaub, Shomir Wilson, Norman Sadeh, Steven M. Bellovin, and Joel Reidenberg. 2017. Automated Analysis of Privacy Requirements for Mobile Apps. In *Proceedings 2017 Network and Distributed System Security Symposium*. Internet Society, San Diego, CA. https://doi.org/10.14722/ndss.2017.23034

[129] Zoom. 2022. Video Conferencing, Cloud Phone, Webinars, Chat, Virtual Events. https://zoom.us/