Is Someone Listening? Audio-Related Privacy Perceptions and Design Recommendations from Guardians, Pragmatists, and Cynics

JULIA C. DUNBAR, University of Washington Information School, United States
EMILY BASCOM, University of Washington Information School, United States
ASHLEY BOONE, PRA Health Sciences and University of California San Diego, Protolab, United States
ALEXIS HINIKER, University of Washington Information School, United States

Smart devices with the capability to record audio can create a trade-off for users between convenience and privacy. To understand how users experience this trade-off, we report on data from 35 interview, focus group, and design workshop participants. Participants' perspectives on smart-device audio privacy clustered into the *pragmatist*, *guardian*, and *cynic* perspectives that have previously been shown to characterize privacy concerns in other domains. These user groups differed along four axes in their audio-related behaviors (for example, guardians alone say they often move away from a microphone when discussing a sensitive topic). Participants surfaced three usage phases that require design consideration with respect to audio privacy: 1) adoption, 2) in-the-moment recording, and 3) downstream use of audio data. We report common design solutions that participants created for each phase (such as indicators showing when an app is recording audio and annotations making clear when an advertisement was selected based on past audio recording).

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; • **Security and privacy → Privacy protections**; **Social aspects of security and privacy**.

Additional Key Words and Phrases: Privacy perceptions, audio recording, smart devices

ACM Reference Format:
Julia C. Dunbar, Emily Bascom, Ashley Boone, and Alexis Hiniker. 2021. Is Someone Listening? Audio-Related Privacy Perceptions and Design Recommendations from Guardians, Pragmatists, and Cynics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 98 (September 2021), 23 pages. https://doi.org/10.1145/3478091

1 INTRODUCTION

People are increasingly surrounded by devices with the ability to record audio. For example, the overwhelming majority of Americans own smartphones; 94% of them say that they frequently keep their phones with them, and 82% say they rarely or never turn these devices off [63]. These powerful computing devices are always equipped with a built-in microphone, a necessity for any phone call. Most tablets and laptops also support high-quality ambient audio recording, and smart speakers, whose primary interaction mechanism involves audio recording, currently have the fastest adoption rate of any technology [73]. This suggests the potential for both users and

Authors' addresses: Julia C. Dunbar, jcd17@uw.edu, University of Washington Information School, 1851 NE Grant Ln, Seattle, WA, United States; Emily Bascom, University of Washington Information School, 1851 NE Grant Ln, Seattle, WA, United States, embascom@uw.edu; Ashley Boone, PRA Health Sciences, Raleigh, NC, University of California San Diego, Protolab, San Diego, CA, United States, aboone06@uw.edu; Alexis Hiniker, University of Washington Information School, 1851 NE Grant Ln, Seattle, WA, United States, alexisr@uw.edu.

2474-9567/2021/9-ART98 $15.00
https://doi.org/10.1145/3478091

their companions to be recorded by devices in their vicinity. Innovations in smart homes and smart cities suggest that the surface area of this exposure is only likely to grow [18].

Activists have raised concerns about audio recording leaving users vulnerable to state agents, technical adversaries, and corporate entities [31, 72]. These fears are not outside the realm of possibility; for example, in isolated incidents, governments have subpoenaed Amazon Echo recordings to investigate crimes [51], and malicious actors have hacked smart baby monitors to listen in on and troll families [14]. Inadvertent data leaks have led companies to share people's private audio recordings with other users [39, 82], and it is a common internal practice at many companies to manually review users' audio data, including recordings of intimate and compromising moments [1]. More mundanely, but at least as consequential, audio data increasingly provides fodder for surveillance capitalism, which introduces users to a host potential harms. By collecting personal data about users' behaviors, companies are able to make decisions about which advertisements to show, strategically predict moments of vulnerability when users will be likely to make purchases, determine what price point a user is likely to tolerate, and more [85].

Prior work has shown that users do not always have accurate mental models of the privacy risks posed by devices with audio-recording capabilities [45], and they regularly grant unnecessary permissions to the systems they use, opening themselves up to additional risk [80]. Further, platforms and devices may not always surface the opportunity to make privacy-related decisions, and for example, a *New York Times* investigation found that more than 250 innocuous-seeming games on the *Google Play* store (e.g., *Pool 3D* by Dumadu Games) access the device's microphone to listen for television programming signals [48]. Without the user's knowledge, these apps can detect the television programming and related advertisements that users are exposed to, collecting data which they can later use themselves or resell to others [48]. Collectively, this past work suggests there are opportunities to design interfaces that better support users' privacy needs with respect to audio recording.

Thus, the goal of this study was to generate design insights to help devices that record audio better support users' privacy. To achieve this aim, we conducted a two-part study, composed of: 1) interviews and focus groups, and 2) design workshops to generate novel design concepts. We found early in our analysis that participants' perspectives regarding audio-data privacy aligned with the existing categories of Schomakers et al. [68]'s framework.

This framework documents that across many different technologies, users can be divided by their privacy behaviors into: *guardians*, who worry a great deal about privacy risks and take action to protect themselves, *cynics*, who worry about privacy risks but do not know what to do about them, and *pragmatists* who worry relatively little about risks and instead focus more on the value technologies offer them. Given that these user groups often require distinct designs to support their needs [68], we examined participants' perspectives and design ideas regarding audio data privacy, differentiated by group. Specifically, we ask the following research questions:

- RQ1: What are the unique concerns and needs of guardians, cynics, and pragmatists with respect to audio data?
- RQ2: What interface designs would best support each of these user types, and collectively, what approaches should designers adopt that will meet the privacy needs of all of these groups?

These groups differed along four distinct axes, and we describe how differences in perspective translate into behaviors. For example, we document guardians' practice of moving away from a smart speaker when discussing sensitive information, and cynics' common experience of attempting to change smartphone audio settings, becoming confused, and abandoning the attempt. We also found that participants' design priorities varied by group. For example, pragmatists' designs blended convenience with privacy, guardians' designs prioritized communicating risk, and cynics created designs that were careful not to assume prior knowledge. Despite these differences, all user groups surfaced three key usage phases that require designed support to protect users' audio-data privacy: 1) adoption or installation, 2) in-the-moment recording, and 3) downstream use of audio data.

Thus, this work offers two main contributions. First, we document the privacy needs of each of these user types in the context of audio recording and present corresponding design concepts. Second, we provide a framework outlining the three key moments in which users can benefit from UI support for protecting their audio-data privacy. This includes making salient when and how audio data is used to shape users' experiences.

2 RELATED WORK

The increase in ubiquitous devices has created a need for new conversations about user privacy. As the number of always-on listening devices have increased, so has the potential for these devices to act as *leaky objects* that "*unintentionally reveal implicit information about individual or collective users*" [34]. Here, we describe prior work investigating privacy concerns in the context of ubiquitous systems that record audio, and synthesize the large body of research modeling users' privacy perspectives.

2.1 Privacy in the Audio Context

Many studies have examined users' perspectives on specific systems with the capacity to record audio. Prior work examining users' perceptions of smart speaker privacy explains that many users are unaware of being recorded and having their recordings permanently stored [49]. Users have varying opinions about the use of this data depending on who is accessing it and what they are using it for [84], personal trust in the company collecting the data [45], and personal interest in the technology.

When people are aware of being audio or video recorded, they consistently express the desire for informed consent [25, 59], and when being recorded by novel technologies, people report feeling more comfortable if they know what the data will be used for [25, 45, 46, 59]. Tabassum et al. investigated user expectations for the future of always-on listening voice assistants and found that participants were able to imagine a vast range of potential services pertaining to conversation but also worry about multi-user settings and the trade-off between privacy and utility [74]. Other work has found that users express cynicism about their lack of control over their audio data [79].

A smaller body of work examines how designers might support users in better understanding and managing their audio data. One investigation examined audio recording in children's toys and concluded designers should implement indicators alerting users when recording is active [50]. An investigation of smart speaker privacy preferences found that smart speakers should be capable of having their microphones disabled via voice commands [45] and in a co-design study of smart homes, participants recommended stand-alone controls for recording devices [83].

2.2 Modeling Users' Privacy Perspectives

A large body of work seeks to model users' general perspectives about their privacy, and numerous classification schemes organize users according to their privacy needs and practices [20, 24, 42, 44, 55, 56, 65, 68–70]. For example, Morton and Sasse demonstrate that users can be clustered according to the types of informational cues they consider important when making privacy decisions, distinguishing, for example, those who look for assurances from reputable companies from those who look to advice from friends and family [55]. Separately, Fortier and Burkell show that Facebook users can be categorized according to their views about the purpose of a Facebook profile, and that these categories predict different privacy needs [24].

Much of the work to theorize and model users' privacy perspectives builds on Alan Westin's program of research. Between 1978 and 2004, Westin published over 30 privacy-related surveys, culminating in the development of a classification scheme for users based on their perspectives about the privacy and security of their personal data. In Westin's final study in this series, he defined three types of users: Privacy Fundamentalists, Privacy Unconcerned, and Privacy Pragmatists [44].

Over time, others came to critique and evolve this taxonomy. For example, Draper explained that much of Westin's depiction of Privacy Pragmatists is built on the idea that these users, "*...have access to the information necessary to evaluate their options and are capable of calculating the costs and benefits...to make optimal choices regarding data privacy*" [19]. Westin's conceptualization of the Privacy Pragmatist has been met with criticism from academics and privacy advocates, who argue that users' behavior frequently contradicts their privacy concerns because they lack sufficient information and are overwhelmed by complexity [7].

The discrepancy between users' reported privacy concerns and actual behavior is known as the *Privacy Paradox* [41, 76]. Prior work documents that this paradox can be explained by cognitive biases that prioritize short-term benefits over the long-term benefits of privacy [6, 7]. Furthermore, the information asymmetries characteristic of privacy decisions can make it difficult for users to fully understand the costs and benefits of privacy and convenience [37]. Users frequently fail to understand what protections they are relinquishing when they grant permissions to third-party applications [22], and apps are often over-privileged and expose users to unnecessary risk [80]. Researchers report that many users describe having privacy concerns but simultaneously express feelings of apathy and futility, which leaves them unwilling to take action to protect their data [32, 76]. Hoffmann et al. dub this behavior *privacy cynicism*, a cognitive coping mechanism that allows users to rationalize their behavior [35]. Thus, categorizing and modeling differences in privacy behaviors requires consideration of these biases and constraints.

Building on Westin's studies and the analysis and criticism that followed, many researchers have explored ways of refining and adapting Westin's classification scheme [20, 42, 56, 65, 68–70]. For example, Dupree et al. sought to improve the predictive utility of the classification scheme, and constructed five personas that reflect a combination of users' motivation and knowledge regarding privacy [21]. The resulting personas, such as the Lazy Expert and the Marginally Concerned, proved more productive for designers than the original Westin categories.

In the current study, we leveraged a categorization scheme developed by Schomakers et al. that further refined Westin's initial model of users' privacy perspectives [68]. They show that users cluster into:

- **Privacy Guardians:** Users that are, "*highly concerned and [take] privacy protective actions.*"
- **Privacy Cynics:** Users who are, "*concerned but [feel] powerless and unable to protect their privacy.*"
- **Privacy Pragmatists:** Users that, "*[show] the least [concern] which they weigh against benefits.*"

We chose to use this categorization scheme to organize our analysis, because it aligned with our first set of emergent themes (see Section 3.1.3) while also addressing key concerns critics raised regarding the original Westin categories. Thus, we leverage this extensive body of prior work to examine users' privacy needs and design priorities regarding audio data through the lens of the known distinctions between pragmatists, guardians, and cynics.

2.3 Designing for Privacy

A large body of work has examined how design decisions can affect users' privacy. To address concerns about privacy, prior work has developed solutions, such as guidance for notification design [52, 71], systems, and tools to protect users [13, 16, 40], government interventions [59, 61], and risk models for IoT devices [33, 36]. Acquisti suggests that "nudges," design decisions that take advantage of various cognitive biases to influence user behavior, can be used to encourage users to make more privacy-sensitive decisions, but questions the ethical implications of this approach [6]. Gambino et al. build on this work by identifying key reasons why users do or do not share their personal information in online settings [26]. Other work has shown that the design of permission notifications changes the likelihood that users will disclose information, [71, 75], and platforms that apply privacy preferences across all apps can significantly decrease the number of decisions users face and the likelihood of sharing sensitive data [13].

In a study investigating voice-activated toys, McReynolds et al. suggest that designers include a change in the visual interface that indicates when a device is recording [50], while Krstulović analyzes the ethical implications of smart home devices constantly "listening" to users [43]. Bugeja et al. recommend that interfaces display risks more intuitively, and offer manipulable functions that give users control over the collection and sharing of data [12]. Other work recommends limiting background-noise recordings [40] and protecting devices from malicious attacks [29]. Finally, other work recommends that the government should regulate privacy protections for users, especially in cases where users feel they do not control their own data [59, 61]. Our work builds on these design recommendations to examine the design directions users feel are most supportive of their privacy needs, specifically within the context of audio data collection.

3 METHODS

We conducted a two-part study to identify user-centered, privacy-conscious designs for devices that record audio data. We first conducted interviews and focus groups to explore users' perspectives on audio recording. We used themes from this formative work to create prompts for subsequent design sessions. Using these prompts, we conducted three in-person and six remote design workshops with a separate set of 14 participants to elicit user-generated design ideas for supporting privacy in the context of audio recording. Throughout this paper, we refer to interview and focus group participants with the notation, "P#" and design session participants with, "DSP#."

3.1 Interviews and Focus Groups

3.1.1 Interviews and Focus Groups—Participants. We recruited 32 participants from across the United States to participate in either in-person semi-structured interviews, remote semi-structured interviews, or in-person semi-structured focus groups. We recruited participants through a mix of channels, including social media, Craigslist bulletin boards, email listservs, groups supporting immigrants with various legal statuses, groups supporting individuals who are blind or low-vision, and an institutionally maintained database of members of the public interested in participating in research studies. Our goal was to assemble a participant pool with a wide range of privacy and security needs. All participants were 18 years of age or older and owned at least one smart device. We did not collect demographic data at the time of the study, but we later reached out to past participants and asked if they would be willing to provide this information. Approximately half of participants responded to this later solicitation, and their data is shown in Table 1.

3.1.2 Interviews and Focus Groups—Materials and Procedures. We developed a semi-structured protocol that probed the participants' perceptions of audio data recording, concerns about privacy, risk-management strategies (if any), usage habits, and other related topics. The protocol for individual interviews was adapted for use in focus groups by adjusting questions, inviting multiple responses, and creating space for participants to build on each other's ideas. The core questions used in individual interviews and adapted for focus groups are listed in Appendix III [1].

Participants were interviewed by a single research team member and were given the option to participate in either a remote or in-person interview. Participants who indicated a preference for a remote interview participated via a video conference call (Google Hangout or Skype) or an audio call. Each focus group was conducted by two research team members in-person at our institution. At the start of each interview or focus group, participants verbally consented to the session being audio recorded. Interview participants received a US$15 gift card to Amazon, and focus group participants received a US$30 gift card to Amazon. Compensation amounts were based

[1] Appendices I, II, and III are online supplementary materials, which can be viewed at the ACM digital library

Table 1. Participant demographic information.

Gender	Women (27%), Men (31%), No response (41%)
Age (years)	18-24 (13%), 25-34 (10%), 35-44 (6%), 45-54 (18%), 55-64 (4%) 65-74 (8%), 75-84 (2%), Prefer not to disclose (2%), No response (37%)
Race	White (31%), Black or African American (6%), Asian (12%), No response (52%)
Ethnicity	Not Hispanic or Latino (41%), Hispanic or Latino (8%), No response (52%)
Vision Impairment	Blind and Low-Vision (27%), General Population (73%)
Highest Level of Education	Less than high school degree (0%), High school degree or equivalent (e.g. GED) (0%), Some college, no degree (2%), Associates degree (8%), Bachelors degree (24%), Graduate degree (14%), Not reported (53%)
Average Household Income	Less than $20,000 (4%), $20,000 to $34,999 (2%), $35,000 to $49,999 (10%), $50,000 to $74,999 (10%), $75,000 to $99,999 (4%), $100,000 to $149,999 (8%), $150,000 to $199,999 (4%), $200,000 or more (4%), Prefer not to disclose (2%), Not reported (53%)

on the length of the session and travel requirements. The study was approved by our university's Institutional Review Board.

3.1.3 Interviews and Focus Groups—Analysis. Audio recordings of all sessions were transcribed and collaboratively analyzed using an inductive-deductive approach [15]. To create the initial codebook, three members of the research team independently read one transcript and inductively created and assigned codes to salient quotes. The team then came together to discuss codes and develop a preliminary codebook. This process was repeated several more times, iterating on the initial codebook. Consistent with best practices in qualitative inquiry [8], we iteratively revised our research questions throughout the analysis process. Our initial code categories distinguished participants as *concerned* or *unconcerned* about audio recording, and early codes categorizing users' perspectives and behaviors included *trust*, *creepiness of recordings*, *control*, *consequences*, and *reactions*. Following Corbin and Strauss [15], we then consulted the literature to examine our emergent themes against existing constructs. We found that these themes aligned with Westin's classification scheme, which categorizes users as Privacy Fundamentalists, Privacy Unconcerned, and Privacy Pragmatists [44]. We then reviewed criticism of this classification scheme and subsequent work that refined these categories [42, 68–70], selecting Schomakers et al [68]'s framework as a useful lens for our data (see Appendix I for a Snip-It of Analysis Based On Schomakers et al. Framework). Once code categories were robust, the research team divided the remaining transcripts for individual coding. The initial codebook contained over one hundred codes, which were refined to yield a hierarchy with fifteen general codes and several sub-codes including: *Trust* with sub-codes such as source trust or socio-political legitimacy and *Knowledge Acquisition* with sub-codes such as passive, stagnant, or proactive. Although we collected data from 32 participants, we reached saturation after reviewing data from 21 interview and focus group participants.

3.2 Design Sessions

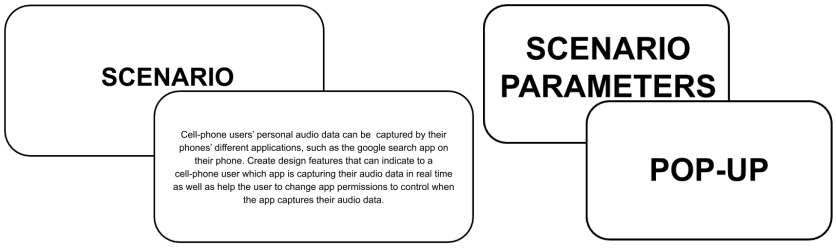
3.2.1 Design Sessions—Participants. We recruited participants through a mix of social media platforms, institutional listservs, groups for people interested in participating in research, and special-interest groups, including those for people who are blind or low-vision. All participants were 18 years of age or older and owned a smartphone or smart speaker. In total, we recruited 19 participants, seven of whom participated individually, and 12 of whom participated together with one or more other participants. Demographic information about participants is shown in Table 1.

3.2.2 Design Sessions—Materials and Procedures. We developed three design session protocols for: 1) an in-person, general-population session, 2) an in-person session for blind or low-vision participants, and 3) a remote session (conducted via Google Hangouts or Skype). As with interviews, participants were given the option to choose between a remote or in-person design session and gave their verbal consent for researchers to audio and video record the session. Participants who completed in-person design sessions received an Amazon gift card worth US$45, and participants who completed remote design sessions were compensated with an Amazon gift card worth US$35. Compensation amounts were determined based on the length of the session and travel requirements.

Using themes from our interviews and focus groups, we created eight "scenario cards" and eight "parameter cards" as prompts for the design sessions. Scenario cards described concrete user scenarios, inspired by the concerns participants expressed during interviews and focus groups. These scenarios focused on misuses of smartphones and cellphones, because these were frequently raised by participants during the interviews and focus groups. Scenario card topics were: (1) microphone access via applications, (2) audio data capture and leakage, (3) real-time notifications of audio data capture and user control over recording, (4) user understanding of audio data recording and its consequences, (5) decision support around audio data, (6) targeted advertising via audio data, (7) user understanding of permission statements, and (8) user monitoring, tracking, and control over social media microphone access.

For example, figure 1a shows a card describing a user scenario where a smartphone application records a user's audio data without their knowledge or consent. The parameters portrayed on the parameter card were provided to help participants scope their ideas. The parameter cards included: (1) light, (2) vibration, (3) noise, (4) pop-ups, (5) text-based disclaimer, (6) icon in the status bar, (7) make/create your own signal, and (8) law/bill/policy. For example, figure 1b) would prompt the participant to design a pop up to address a scenario about applications accessing the microphone.

In-person design sessions began with introductions and pre-task discussion (25 mins), followed by individual and group design tasks (35 mins), post-task discussion (20 mins), and wrap-up (5 mins), for a total of 85-90 minutes. For the design tasks, each participant was given markers, a large sheet of paper, one random scenario card, and one random scenario parameter card. Participants individually sketched solutions to their given scenario and scenario parameter cards and then repeated the exercise in pairs with another participant. After completing the design-task, each pair presented their favorite solution and the researcher facilitated a conversation about their designs. These procedures are outlined in more detail in Appendix II.



(a) Scenario Card (b) Parameter Card

Fig. 1. Examples of a Scenario Card and a Parameter Card, which would be randomly assigned to participants during design sessions.

Remote design sessions followed the same general format but were modified to accommodate the online context. All remote design sessions were conducted one-on-one with a research team member and the participant individually completed two to four rounds of the design task. For each round, the scenario and scenario parameter were randomly selected and read aloud to the participant. To accommodate blind and low-vision participants, we provided these participants with three pdf or word documents in advance: the design session schedule, the scenarios, and the scenario parameters. Participants used a screen reader for the design session as needed. Participants were given the option to either verbally communicate their design solutions or document them on paper and send them to the facilitator at the end of the session.

3.2.3 Design Sessions—Analysis. After transcribing design session audio recordings, two members of the research team conducted an initial analysis of the transcripts and design artifacts (see examples of participant design artifacts in figures 3a, 4a, and 5a). During this analysis, researchers identified preliminary themes and added them to the codebook created during the analysis of interview and focus group data. The research team then categorized design session participants according to Schomakers et al. [68]'s framework (see Appendix I for a Snip-It of Analysis Based On Schomakers et al. Framework). Although design participants did not answer formal interview questions as interviewees and focus group participants did, the research team was able to categorize design session participants by focusing on participants' responses during the pre-task discussion and conversations participants had surrounding their perspectives on audio data recording while completing the design tasks.

After participants were categorized via Schomakers et al. [68]'s framework, two research team members then re-examined participants' design ideas by group (pragmatists, guardian, or cynic) and through collaborative discussion and iterative coding over several weeks, they identified design themes within each group. Although 19 individuals participated in the design sessions, saturation was reached after analyzing data from 14 participants. Three members of the research team then conducted a secondary iterative analysis of design ideas through weekly meetings over several months, in which designs across all three categories were looked at holistically.

Finally, the research team created wireframes using a prototoyping tool called Figma [5] (see Figures 3b, 4b, and 5b) through collaborative discussion and interpretation over several weeks as a way to make sense of participant design ideas.

4 RESULTS

4.1 Perspectives on Audio Data Collection

We found that participants' perspectives aligned with the framework developed by Schomakers et al. [68]. Specifically, their perspectives reflected the attitudes of: guardians (n = 10), pragmatists (n = 15), and cynics (n = 10). We further found that participants' behaviors differed by group, as described in detail below. Figure 2 summarizes some of these themes, showing four ways in which group behavior differed. For example, a majority of guardians (represented by "G" and a red circle in Figure 2) said they physically move away from smart devices or lower the volume of their voice when discussing something sensitive. In contrast, cynics (represented by "C" and a yellow circle) reported occasionally but inconsistently engaging in this behavior. Pragmatists (represented by "P" and a blue circle) reported rarely regulating their distance or voice when around a microphone. We elaborate on these behavioral differences below.

4.1.1 Perspectives of Audio Pragmatists. Fifteen participants (43%) expressed views that were dominantly reflective of the pragmatist perspective. As described above, privacy pragmatists are those who have few privacy concerns and show a willingness to trade privacy risk for high-value experiences. Their trust in large companies and a high degree of technology self-efficacy help them to feel comfortable adopting technologies that may come with some risk [68].

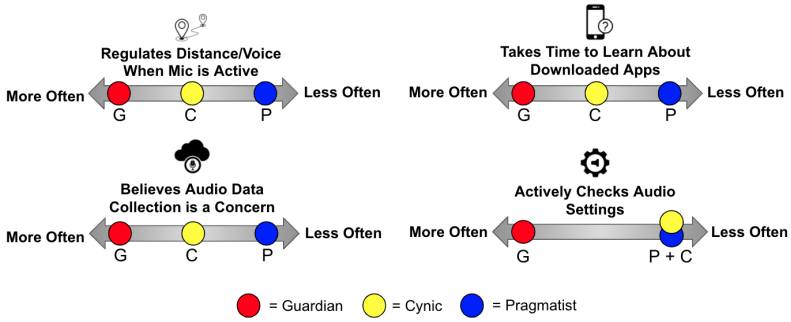


Fig. 2. The behaviors and perspectives of the three user groups (guardians, cynics, and pragmatists) and how they differ along four different axes: 1) regulating their physical distance/voice when a microphone is active, 2) belief that audio data collection is a concern, 3) takes time to learn about downloaded apps, and 4) actively checks audio settings.

Audio-data pragmatists explained that, for them, the benefits of enabling microphone features compensate for potential risks. P16 explained that, even after learning about risks posed by audio data collection, they are not interested in changing their daily behaviors to mitigate risk, "*'cause I think right now the pros outweigh any issues. It's just easier for me to really contact people and use my technology more [by using the microphone], especially when I'm driving.*" P20 emphasized the value of audio features, saying, *"If I am doing something else and I need to… call someone I do the, 'Okay Google'… They [Google] need to use the microphone so you just give it access because there's no other way to actually use the feature.*" Pragmatists explained that they avoid audio-recording features to protect their privacy only when these features offer so little value that there is no reason to use them. P22 explained they would disable the microphone only, "*if it didn't impact my life too much*," and P25 elaborated, "*I never went in and tried to turn anything off [in the settings] 'cause I use it all the time! [laughs].*" This focus on the value of audio data collection was characteristic of pragmatists across interviews, focus groups, and design sessions.

Pragmatists explained that their focus on the value, rather than the risk, of microphone-enabled features stems in part from their faith in large companies. They said things like, "*if it's coming straight from Google, that's fine… if Google was using data [inappropriately], they'd be pretty screwed, so in that sense I don't think the company would take the risk in using the Google Home to farm my data and steal my identity*" (P1). Similarly, DSP2 explained, "*They may be monitoring occasionally, but I kind of just have to trust the app that they're only using the microphone when I want them to use the microphone.*" This innate trust in large online companies like Google, Amazon, and Apple gave pragmatists the confidence to use audio-enabled technologies without taking action to safeguard their data. Participants explained that these large entities are regulated and subject to both government oversight and public opinion, explaining that this accountability gives them faith that companies will not use audio data in problematic ways.

Because of this trust, pragmatists stated that they were not concerned by the idea of devices with always-on listening, even if it extended beyond well-defined usage scenarios. P16 explained that their device listens in as they go about their daily life, but this invasion is a not a threat: "*I'm sure it's listening to me all the time. [laughs] For example, cause I do use the 'Okay Google' a lot, so it's always listening to me. Like sometimes I'll be in a meeting and I'll [hear]—[beeping impression] 'I did not recognize that voice'—or something like that… it's listening to what I'm saying.*" Across this group, we found that participants occasionally invested in learning about audio features and settings but typically did not engage in protective actions or express concerns about data collection.

4.1.2 Privacy Guardians. Ten (29%) participants expressed views that were primarily reflective of a guardian perspective. These participants described regularly taking steps to safeguard their audio data or prevent it from being collected in the first place. For example, DSP8 described their own behavior saying, "*The microphone, I don't have it by default in the apps, only when I want to use the app…I don't have anything like that enabled by default.*" DSP10 goes to even greater lengths to prevent unnecessary audio data collection; like other guardians in our study, they modify what they say in front of their smart devices: "*Once I have my phone with me I'm very cautious of what I say or what I talk about…if I don't want to be quoted, I just keep quiet.*" Many other guardians told us they avoid technology altogether if they feel it might be collecting audio data invasively. For example, DSP2 explained, " *I will look at the permissions…that app has no reason to use this data and so, I just flat out won't download it.*"

In addition to avoiding systems and shutting off features, guardians told us that they work to protect themselves from audio data collection by learning about the technologies they use. Participants said things like:

> "*If I hear about a particular app, of course I go to Google and do a Google Search–that is [step] one: to read about it. Then, two, you always find, you know, user reviews and you see five star, four star, three star, two star. I go after the worst of the comments.*" (P24)

Guardians commented on the importance of learning about the audio recording behaviors of the technologies they use, their interest in this learning process, and the fact that they wait to engage with audio-recording features until they have had a chance to do this due diligence. For example, DSP5 explained:

> "*I have no apps turned on, partly because I'm really hesitant for all the little things they want to know and want permission for, but I can tell that what I really need is a private session with somebody to learn how to turn things on and off…I would find that very interesting.*"

These participants described this background research as an essential prerequisite to using audio features.

4.1.3 Privacy Cynics. The remaining 10 participants (29%) expressed views that were predominantly reflective of a privacy cynic. Cynics described accepting default permissions and settings, expressed a lack of understanding of how and when their devices collect audio data, and described uncertainty about how they might learn more. These users described their lack of understanding of audio data collection by saying things like, "*I'm just not sure if that data is stored on the phone or…if it's out in the air in like their databases or what not*" (P13), and "*I can't figure out how to turn it off*" (DSP7).

These participants lacked pragmatists' trust that technology companies will protect them, but they also lacked guardians' self-efficacy and confidence that they can protect themselves. As a result, they described floundering: either attempting to protect themselves through actions they do not fully understand or giving up and doing nothing. They described wanting to prevent audio recording but not knowing what to do about it, saying things like, "*You know, I've tried to [change the audio recording settings], especially when I first got the phone many years ago, and I found it screwed things up too much, and I don't know if it really made any difference*" (P14). Similarly, DSP12 described wanting to adjust their microphone settings, but not knowing how to do so or how to learn more: "*I don't know how you turn off the microphone…I looked and I don't know where there's a place to.*" They explained that, unlike guardians, they do not feel confident in their ability to master the technologies they use or take steps to educate themselves about managing their audio data. Yet unlike pragmatists, who said they choose not to take action to safeguard their data because they trust the companies who collect it, cynics told us that they do not take action because they are not sure how.

Cynics also said they lack confidence in their understanding of the risks posed by audio data collection, saying things like, "*I don't have a great understanding of how my information could be used against me*" (P15). This lack of clarity about how to protect themselves and what they might be protecting themselves from left cynics with a sense of resignation and disengagement, leaving devices in their default state to collect audio data unimpeded.

For example when P3 told us, "*I haven't turned anything [recording devices] off*," they further explained that this is because, "*I don't know what kind of stuff or risks there can be, so I'm like, 'eh, what's the point of thinking about it?'*"

4.2 Design Ideas from Pragmatists, Guardians, and Cynics

This section discusses design ideas from the design session participants ($n = 14$) and from two interview participants (P27 and P28) who inadvertently discussed design ideas during their interviews. We found the three different perspectives (pragmatist, guardian, and cynic) each led to unique design priorities. Pragmatists' designs reflected a dual interest in protecting themselves from invasive recording and giving themselves easy access to recording features they value. Guardians created design solutions that allowed users to look up granular information on demand. And cynics created designs that communicated security risks clearly and did not assume prior knowledge. Despite their differences in design needs, all three groups independently created design concepts that addressed three common scenarios: 1) installation or adoption of technologies that record audio, 2) in-the-moment recording, and 3) downstream use of users' audio data after it has been collected. Here, we describe the designs they proposed in each of these three contexts.

4.2.1 Installation-Time Design Concepts. Participants of all types independently created designs that addressed the way smart devices inform a user about audio recording and request permission when the user first adopts the technology. Many participants talked about end-user license agreements (EULAs) and proposed designs to make EULAs more effective in supporting privacy needs. These redesigns universally proposed simpler interfaces that were easier to understand, with participants expressing frustration with the status quo. They said things like, "*How about making them one page instead of 35?*" (DSP9). Yet despite users' common interest in a simplified experience, each user group (pragmatists, guardians, and cynics) chose to emphasize and make salient slightly different dimensions of the EULA and the process of informing users about the details of the system.

Pragmatists wanted EULAs to highlight recording features so they could make quick adoption decisions based on the value of the feature and the circumstances under which recording would occur. For example, P27 explained that, "*if it [the EULA] highlighted the main points, or the main risks and benefits, I would definitely download it [the app]*" (see Figure 3b). Similarly, P28 designed a EULA with "*filters*" to give the user the ability to focus quickly on the specific feature or permission of interest. And DSP1 explained that at install-time, "*everything…should always be that you have to opt in.*" In these instances, pragmatists described wanting straightforward interfaces that make it easy to quickly grasp risks and benefits.

Guardians also wanted more approachable EULAs, but they emphasized the importance of highlighting risks and helping users protect themselves. In their ideation, they suggested changes like, "*What if the danger to a user is highlighted and enlarged on the screen? You know the magnifying application that allows you to magnify it? If the program in the agreement were to say, 'Are you sure agree to this?'*" (DSP3) (see Figure 3b). Similarly, DSP5 describes adding icons to make risks more salient: "*…if I open up the agreement, there are icons that tell me what's important.*" (see Figure 3b).

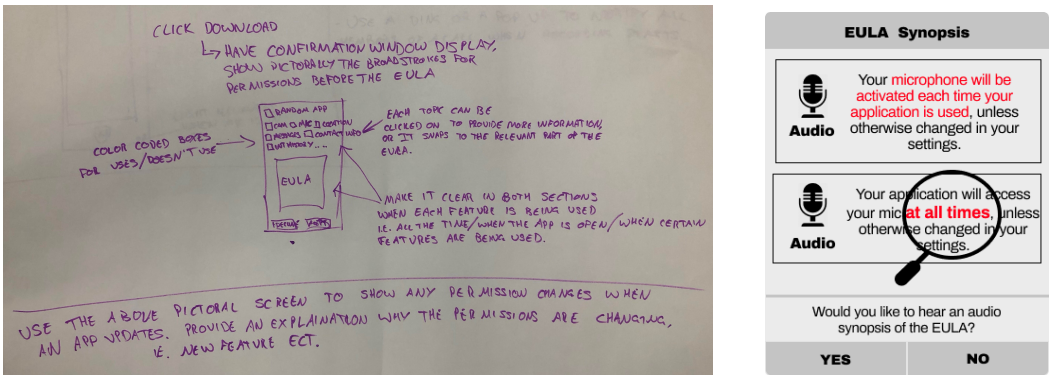
Cynics were even more adamant than other user groups about the importance of simplifying the EULAs. In addition to the overview of features valued by pragmatists and risks valued by guardians, these users explained that the design of any interface component in which users consent to recording should: include confirmations, avoid assuming that the user has a strong grasp of how the system works, and make every effort to ensure the user truly understands what they are agreeing to. For example, DSP7 described a EULA interface they designed, saying:

> "*You get a pop-up box that says, 'This app would like to use your audio services and would you agree to that or not?' If you say yes, then you get a second box that says, 'This app will now have permission to record you at any given time,' or, 'It will record you when the app is in use,' or whatever specific situation*

that they want to have the access to the audio… once you agree to it, then you have to double-agree to it, so it says specifically, 'Are you sure that you want to?'"

As shown in this example, cynics' solutions not only requested consent from users but also reflected back to them what they had consented to.

Across user types, many participants requested systems that would play EULAs in audio format (see Figure 3b). They said things like, "*I would want to have the agreement spoken to me out of the audio*" (DSP4), "*an audio version of the agreement would be awesome*" (DSP12), and "*they need to have like an audio synopsis accompanying the text… like a downloadable MP3 or a playable MP3 which just gives you the bullet points*" (DSP9). These users and others who requested audio EULAs were blind or low-vision participants. This concern was orthogonal to the privacy orientations we identified and suggests an unmet need among users who cannot access content visually.



(a) Participant Design (b) Wireframe

Fig. 3. Example wireframe capturing common themes in participants' design solutions about the beginning of the consent process.

Example Wireframe 1: An Installation-Time Interface. Drawing on these insights, we created the wireframe shown in figure 3b, this composite design reflects a combination of themes in participants' ideas for improving installation and adoption. The first design theme came from pragmatists, who wanted easier EULA navigation. We addressed this theme by adding simple audio icons to make the topic visually obvious. The second theme came from guardians' interest in raising users' awareness of risks. In response, our design highlights the most prominent risks in red, and includes a magnifying glass that emphasizes risk-related text. The third design theme was inspired by cynics' focus on repetition of important points. This was incorporated by repeating the same or similar message twice. Lastly, we added an option to listen to an audio synopsis of the summary points of the EULA.

4.2.2 Design Concepts Regarding In-The-Moment Recording. Across user groups, participants also created designs that notified the user in real-time whenever audio was being recorded. These ideas ranged from raised braille notifications to added color masks to persistent notifications in the status bar and notification shade. One pragmatist participant shared a design in which the start of a recording was preceded by a sound that is, '*as distinctive as possible. So, it's like three sharp beeps, or your favorite songs so you can customize it… or clapping noise, stuff like that*" (DSP11). Another pragmatist described wanting, *"an icon, maybe in the top right–like the battery or Wifi area… something basic like a blank circle or square or something*" (DSP1) (see Figure 4b). Pragmatists explained that despite their interest in being notified about recording, they would want these notifications to fit

the constraints of daily life. For example, they added limitations to their own designs to ensure that they were not disruptive, saying things like, "*[The notifications should] turn off during some time of the day. I feel like during 1am - 6am*" (P28).

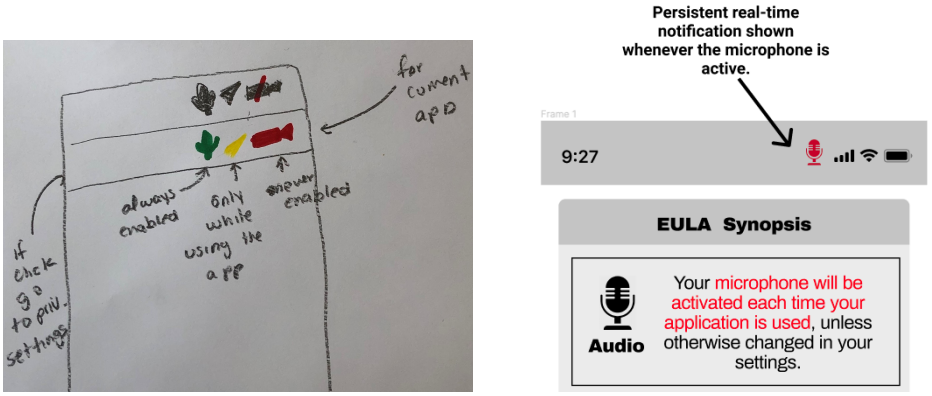
Guardians also designed real-time notifications alerting them to microphone activation. These designs were more detailed and provided more information to the user, frequently indicating not just that the microphone was recording, but specifically what experience had triggered it, which third-party experience was behind the recording, and what their motivation was. For smartphone recording, guardians often wanted indicators that reported which app was accessing the microphone. For example, DSP3 said,"*if an application is following you, that application should light up on your phone. You should be able to see that Facebook is following you it should have a little light around it.*" DSP9 also wanted app-level cues to signal who exactly was accessing their microphone, suggesting, "*the app itself [should] switch colors when it was in listening or watching mode.*"

Guardians also designed features that allowed users to preemptively avoid recording before it begins, consistent with the avoidance behaviors unique to this user group. For example, DSP14 suggests, "*If you're about to start talking about something you don't want your microphone to pick up, maybe design a way where it can tell that you're starting to talk about stuff like that and it will [ask], 'Do you want us to turn off your microphone for you?'*" In these design solutions, devices worked together with users to find ways to prevent invasive recording from ever beginning.

Cynics also designed audio and visual feedback to notify the user when recording begins. These users prioritized making their designs simple and easy to use, saying things like, "*the light simply could indicate that either recording is on or it's off—that the microphone is on or off—so you don't have to go through all of that [work to understand and adjust permissions]*" (DSP4). Cynics also prioritized ensuring the user understands what is happening, saying things like:

> "*[A] pop up should allow the mic to be on or to be off. It doesn't have to do it continuously,... but when you close [the app] and open it up again it should pop up again... That would give me a lot of information, because I don't know how my apps, how many of them might have mic capability.*" (DSP12)

This kind of double confirmation was characteristic of design concepts created by cynics.



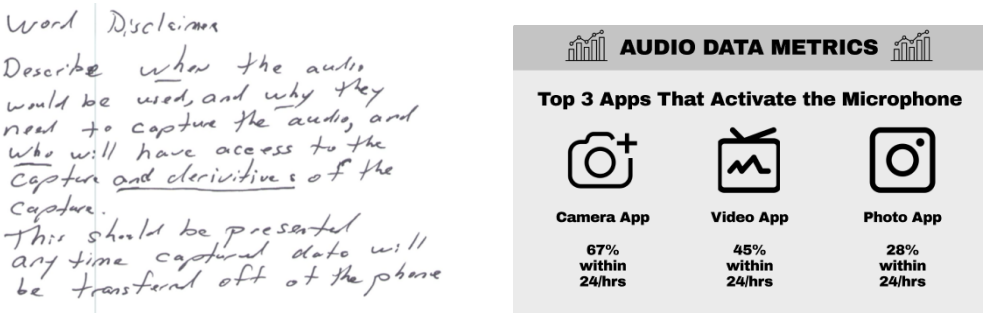
(a) Participant Design (b) Wireframe

Fig. 4. Example wireframe of a participant's design solution for a real-time notification that depicts a microphone in the status bar of a phone, which turns red when activated. Multiple participants came up with similar persistent real-time notification ideas.

Example Wireframe 2: In-the-Moment Recording. The research team created the wireframe shown in figure 4b, which was inspired by the in-the-moment notification designs of each participant group. Drawing on common ideas from participants, this design shows a red microphone icon in the phone's status bar any time the microphone is active. Pulling down the notification shade could provide further detail, such as indicating the app responsible for recording and the ability to stop recording.

4.2.3 Design Concepts Regarding Downstream Use of Users' Audio Data. Participants also expressed a need for designs that show users the downstream effects of audio data being collected. The design ideas and values surfaced were consistent across user types, as opposed to the phases described above where we observed nuanced differences in the solutions created by pragmatists, guardians, and cynics. Participants explained that, in addition to being notified of audio recording when it happens, they would also like to be notified when their audio data is used. For example, DSP11 described a scenario where a hypothetical user was exposed to an advertisement for dog food, selected because the user had previously discussed wanting to purchase dog food. The participant created a design that "*has a blocking feature, so once the company targets these previous conversation like dog food ad or a shopping ad and you don't have to hear it anymore then you can block these through the pop up.*" In this design concept, the user is shown concretely how their audio data has been used and given the option to undo that action. Similarly, DSP14 designed a system to notify the user when it detects audio that it will use to inform future decisions. DSP14 explained that "*During a phone call or something like that, you'll get a little vibration during the call and it'll tell you that something picked up a keyword that they'll use.*" In this way, the system not only makes salient the fact that it is recording; it specifically highlights the moments of data collection that are of consequence.

Other participants suggested analytic or dashboard views that would make it easier to review what had been recorded in the past and who had collected this data as a way for users to assess risks retrospectively (see Figure 5b). For example, DSP9 described wanting "*a section in the app that has like a visual representation, maybe like a table of how much time your app has been spending listening to you or when the last time it was listening to you or if it's currently listening to you, kind of an ongoing, like, visual statistic log of that information.*"



(a) Participant Description (b) Wireframe

Fig. 5. Example wireframe inspired by multiple participant's written design ideas and verbal design descriptions.

Example Wireframe 3: Downstream Use of Users' Audio Data. We created the wireframe shown in Figure 5b as an illustration of the audio-data dashboard that participants described. The dashboard displays a visual overview of applications that have collected audio data, with the ability to drill into any one application's data in more detail.

5 DISCUSSION

We found that participants' audio data perceptions clustered into the categories defined by Schomakers et al. [68]'s framework. Although participants shared many of the same cross-cutting concerns, these user groups also reflected distinct privacy needs. Pragmatists valued designs that foregrounded a feature's benefits in addition to its risks, guardians wanted protections that come *before* any recording occurs, and cynics wanted systems that would provide double-checks to reduce misunderstandings. These differences in priorities translated into slight differences in the design solutions participants invented, such as install-time dialogues that summarize the trade-offs a user is making, intelligent microphones that turn themselves off when users might be discussing something sensitive, and pop-ups that paraphrase and confirm a user's decision before acting on it.

5.1 Protecting Users Across the Recording Life Cycle

Participants from each category and across all design sessions invented designs that spanned the life cycle of a recording event and highlighted three phases of risk (see Figure 6). This suggests that protecting users' privacy with respect to audio data collection requires intervention at each of these points in time. Users provided concrete suggestions for supporting privacy in each phase.

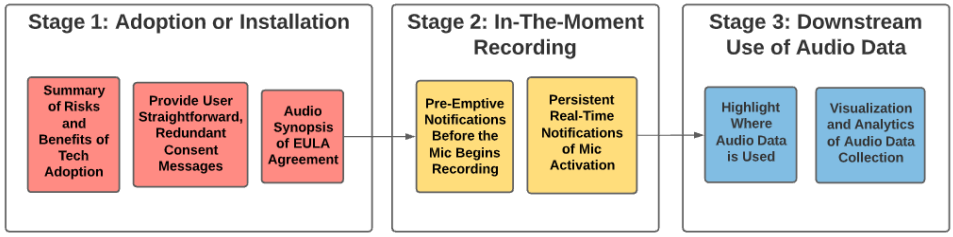


Fig. 6. User Experience Pipeline and Interventions Audio Model

5.1.1 Stage 1: Adoption or Installation. Participants called for simplified EULAs that support informed choices when the user first adopts an experience that will access the microphone. Their ideas suggest the following guiding principles:

- **Support Cost-Benefit Analysis:** Provide users with a distilled summary of the main risks and benefits of the experience they are about to adopt. By surfacing the benefits, designers will support the needs of pragmatists as they make calculated decisions, and by making the most severe risks salient, designers meet the needs of guardians.
- **Strive for True Consent:** Provide straightforward messages, potentially with redundancy. All users designed install-time consent to be clearer and simpler, with cynics particularly interested in redundant designs that would help ensure the user knows what they are agreeing to.
- **Provide an Audio Synopsis of EULAs:** Improve the accessibility of informed consent by making these distilled summaries available via audio.

All participant groups highlighted the importance of this stage. The support they needed varied slightly, for example, with pragmatists emphasizing their interest in quickly understanding the risk-benefit trade-off and cynics wanting interfaces to ensure the user really understands what they are consenting to. But none of these needs were in conflict, and a thoughtfully designed installation-time interface could meet these needs simultaneously. Unfortunately, participants' frustration with EULAs is not new, and many prior studies have

documented the shortcomings of companies' consent practices [10, 30, 38, 57, 58, 67, 77]. Prior work shows that users are conditioned to accept EULAs uncritically [10]–even to the point of consenting to install spyware [30]–and that the literacy burden of existing EULAs exceeds the comprehension level of many adults [47]. Researchers have also created novel designs demonstrating how companies might improve the adoption experience and increase the transparency of EULAs, such as automated systems for surfacing key take-aways [57, 58] and animated sketches modeling how the technology works [54]. Despite this existing design guidance, EULAs and other consent agreements remain opaque and intractable for users. Our findings indicate that improving the design of EULA interfaces is essential to empower users to make privacy-conscious decisions about their audio data, suggesting the need for regulation that places demands on the design of installation UI.

5.1.2 Stage 2: In-the-Moment Recording. Participants from each category came up with even more designs that intervened while the microphone records audio. Their collective designs suggest the following design principles:

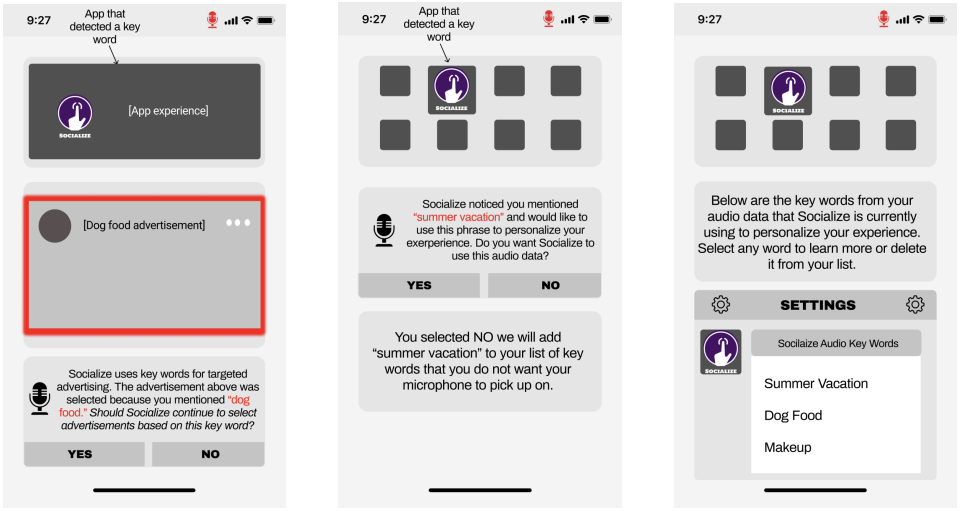
- **Provide Persistent Notifications:** Create a persistent indicators that makes it easy for the user to recognize when recording is live and keep this indicator salient for the duration of the recording.(see Figure 4b)
- **Provide Preemptive Notifications:** Create preemptive design solutions to inform users before the microphone starts recording to ensure notifications raise users' awareness before data collection begins, a principle consistently surfaced by guardians.
- **Connect Recording Indicators to their Source:** Help users understand who is recording audio and why.

All participants emphasized the need for interfaces to make active recording obvious to the user. In some instances, existing platforms have already begun to provide such support; for example, in a recent update, iOS added an orange icon showing when the microphone is actively recording [66], and third parties have created apps to mimic this experience on Android [4]. The existence of these features is a promising indicator that companies are interested in designing to increase the transparency of audio recording, despite the fact that collecting audio data secretively and invasively may serve corporate interests. Participants' design ideas both align with current commerical designs (e.g., valuing persistent indicators) and offer richer and more nuanced versions than those that are available today.

5.1.3 Stage 3: Downstream Use of Audio Data. Participants of all types suggested design ideas for surfacing the effects of past audio recording to raise users' awareness of general recording practices and their consequences. All user groups expressed a need for designs to show users the effects of audio data being collected, and they created designs to support users' learning and conceptualization about what is happening with their audio data. The following collective design suggestions emphasize possible points of intervention to help users with their understanding of downstream use of audio data:

- **Show Results:** Make visible the ways in which interface elements are influenced by audio data, such as highlighting advertisements selected as a result of spoken words or adding icons to visuals that reflect audio history (see Figure 7a).
- **Provide Notifications for Downstream Effects:** Alert users to moments when real-time processing detects audio of interest to the system that might be used later. (see Figures 7b and 7c).
- **Present Data Summaries:** Provide visualizations and analytics about audio data collection. For platforms that give recording access to third-parties, create glanceable interfaces that allow users to view recording statistics for each third-party experience (see Figure 5b).

Participants told us that they lacked a crisp understanding of what happened to their audio data after it was collected, consistent with prior work showing that people often struggle to anticipate the concrete threats posed by an interface, even when they know it is manipulative [11]. They designed solutions in which products help them understand what audio is collected, when audio data is used, and how moments of recording connect to changes to the interface. We created three example wireframes inspired by these design ideas: Figure 7a illustrates



(a) "Show Results." In this example, the app connects an advertisement to the audio data that led to its selection

(b) "Provide Notifications." In this example, the app alerts the user (and allows them to intervene) when it is about to file away a piece of audio data for future use

(c) "Present Data Summaries." In this example, the user reviews the pieces of audio data that the app is currently using to shape the experience

Fig. 7. Wireframes that depict examples of the three downstream use of audio data suggestions.

participants' interest in UI to reflect which app records what data. Figure 7b illustrates participants' interest in seeing how real-time recordings are used by the system (with the potential for the user to intervene). And Figure 7c reflects participants' interest having a mechanism for user-driven review of the audio data that has been collected and how it is currently leveraged to shape their experience (again, with the potential for the user to intervene).

These design concepts illustrate just a few mechanisms by which designers might surface the concrete effects of audio recording. Relative to the design of EULAs (which has been studied extensively) and the design of real-time notifications of recording (which has begun to emerge as a design pattern in commercial experiences), the design of UI to communicate the downstream effects is relatively under-explored. Future work to more deeply solicit users' input into the design of such interfaces would be useful for informing both design decisions and policy.

Regardless of their privacy orientation as a pragmatist, guardian, or cynic, all participants made clear that they need support from designers at each of these three phases of the recording lifecycle. Only by addressing all three can designers protect users from *privacy leakage* in the context of audio recording [62]. Audio data is increasingly collected by artifacts that traditionally would pose no data collection risks, including objects like buildings [64] and dolls [9]. We encourage designers of any system recording audio—whether embedded in the walls of a room or a child's toy—to consider: 1) how to provide support for each of the three phases outlined above, and 2) how they might align their design with the principles participants surfaced for each stage. Privacy breaches in this space are frequent (e.g., [2, 17, 60, 78]) and government entities are increasingly adopting user data privacy regulations (e.g., [23, 28]), forcing companies to make changes to the way they collect and use users'

information. Based on our findings, we propose that surfacing the downstream effects of audio data collection should be a priority for application designers and a component of privacy legislation.

Finally, we note with optimism that the design ideas generated by each user group are largely not in conflict, suggesting it is possible for designers to create composite interfaces that support the needs of all three user groups. However, this may not always be the case: the repetition that cynics proposed could be seen as a nuisance by pragmatists, and foregrounding value propositions–as proposed by pragmatists–could be viewed by guardians as inappropriate. We see potential for designers to minimize these possible value tensions [53] and to meet the needs of all user groups simultaneously. However, doing so may require proactive consideration of potential collisions and a commitment to the principles of universal design [3]. Future work remains to further examine this possible design challenge.

5.2 User-Driven and System-Driven Behavior

We found that participants often felt forced into making choices regarding audio data collection that did not reflect their desired course of action. Participants who were concerned about their audio data being recorded (mainly those categorized as guardians or cynics) reported taking different actions to protect themselves from invasive data collection. These participants described moving away from smartphones to have sensitive conversations and leaving the room to discuss certain topics when they feared their smart speaker might listen in. Other users described encountering apps they wanted to download but could not because the app's permissions requests were too intrusive. These behaviors reflect ways in which users deviate from their preferred course of action (i.e., downloading an app they wish they could use, or continuing a conversation in the room they are already in) to conform to the demands of the technology.

This struggle is not unique to invasive audio recording or to questions of data privacy. For example, Gardner and Davis describe young people's use of social and communication technology as alternatively *app enabled* or *app dependent* [27]. App enabled usage empowers the user to engage in behavior they prefer to pursue, which coincides with pragmatists' desires to not have the technology get in the way of their daily lives, such as sending a Snapchat snap out of a desire to connect with a close friend. In contrast, app dependent usage chases the demands of the system, perhaps sending a snap to that same friend, because the app dictates that this is the only way to maintain a snapstreak.

Separately, Wobbrock et al. define the *ability-based design* paradigm in contrast to prevailing trends that demand users of varying abilities and disabilities adapt themselves to the needs of a system [81]. Ability-based design calls for designers to adopt a fundamental orientation of holding systems, rather than users, accountable for poor performance and recommends designing interfaces that adapt to make themselves useful and usable to people of all cognitive and motor abilities.

The app dependence documented by Gardner and Davis and the status quo that ability-based design rejects are two of many examples of interface patterns that undermine users' autonomous engagement with the world. Although technologies that enact invasive recording, promote dependence and addiction, or present inaccessible interfaces may differ dramatically in many respects, they all push users to engage in behaviors that compensate for shortcomings of the system. Here, we differentiate between *system-driven behavior* and *user-driven behavior* to highlight the distinguishing feature of this pervasive and cross-cutting design problem. We find that invasive audio recording has the potential not only to compromise users' privacy but also to push users toward system-driven behavior in which the technology, rather than the user, is the guiding actor. By addressing users' concerns about audio recording, designers would not only safeguard user data, they would also reinstate users' power over their own behavior, with the freedom to conduct conversations in any room they please.

The forced choice between privacy and access is consistent with the standing debate of the privacy paradox, where past work has shown that the coercive nature of design, rather than a lack of interest in protecting

one's own data, drives users to compromise their privacy [6, 37, 76, 79]. The sentiments of our participants are consistent with those of many users of many systems. The all-or-nothing decisions that designers create for their users lead some users to disengage entirely (e.g., choosing not to download apps of interest), and other users to give up something they value (e.g., protection from intrusive audio).

5.3 Limitations and Future Work

Although we recruited a diverse mix of participants, there are numerous ways in which users might have unique privacy needs or vulnerabilities with respect to audio-recording that are not reflected here. For example, survivors of intimate partner violence and individuals who use shared devices would likely illuminate insights we do not uncover. Also, all of our participants currently reside in North America and may exhibit a general position towards privacy that may not be comparable to other parts of the world. Our findings are also limited in that they rely on self-report and speculation about future use, and as noted in the literature surrounding the privacy paradox, there is often a mismatch between users' reported attitudes about privacy and their actions [41, 76]. There are also many ways in which users' privacy needs can be understood and categorized, and Schomakers et al.'s framework is just one of many possible organizing mechanisms. Future work drawing on other lenses would likely reveal additional needs and design directions. Although our participants present a number of novel design solutions, we do not assess them here, and future work remains to implement and evaluate these concepts. Future work to evaluate design approaches for highlighting the downstream effects of collecting audio data would be a particularly valuable complement to the results presented here. Finally, we clustered design participants into Guardians, Pragmatists, and Cynics based on their conversation and contributions during the design session without probing their perspectives through a structured protocol.

6 CONCLUSION

In this paper, we report on a two-part study with 35 participants to elicit design ideas for making devices that record audio data sensitive to users' privacy. We find that pragmatist, guardian, and cynic perspectives are all well-represented with respect to audio recording, and we find that each group has distinct privacy needs in this context. Despite their behavioral and attitudinal differences, all three groups surfaced the same three usage phrases as moments that require design consideration with respect to audio recording: adoption of a new technology, in-the-moment recording, and downstream moments when audio data is used. We contribute a set of principles for supporting all three types of users across these three phases of the audio recording lifecycle.

ACKNOWLEDGMENTS

Thanks to Dylan Hardman, Estelle Jiang, Leanne Liu, Allen Shi, and Yuxing Wu for their help with this research. This work was funded by the University of Washington Royalty Research Fund.

REFERENCES

[1] 2019. Improving Siri's privacy protections. https://www.apple.com/newsroom/2019/08/improving-siris-privacy-protections/
[2] 2019. Your Data Is Shared and Sold…What's Being Done About It? *Knowledge@Wharton* (Oct 2019). https://knowledge.wharton.upenn.edu/article/data-shared-sold-whats-done/
[3] 2020. What is Universal Design. http://universaldesign.ie/What-is-Universal-Design/
[4] 2021. Access Dots - Android 12/iOS 14 privacy indicators - Apps on Google Play. https://play.google.com/store/apps/details?id=you.in.spark.access.dots
[5] 2021. *Figma*. https://www.figma.com/
[6] Alessandro Acquisti, Idris Adjerid, Rebecca Balebako, Laura Brandimarte, Lorrie Faith Cranor, Saranga Komanduri, Pedro Giovanni Leon, Norman Sadeh, Florian Schaub, Manya Sleeper, et al. 2017. Nudges for privacy and security: Understanding and assisting users' choices online. *ACM Computing Surveys (CSUR)* 50, 3 (2017), 44.

[7] Alessandro Acquisti and Jens Grossklags. 2005. Privacy and rationality in individual decision making. *IEEE security & privacy* 3, 1 (2005), 26–33.
[8] Jane Agee. 2009. Developing qualitative research questions: a reflective process. *International journal of qualitative studies in education* 22, 4 (2009), 431–447.
[9] BBC. [n.d.]. German parents told to destroy Cayla dolls over hacking fears. https://www.bbc.com/news/world-europe-39002142#:~:text=An%20official%20watchdog%20in%20Germany,Bundesnetzagentur)%2C%20which%20oversees%20telecommunications.
[10] Rainer Böhme and Stefan Köpsell. 2010. Trained to Accept? A Field Experiment on Consent Dialogs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 2403–2406. https://doi.org/10.1145/1753326.1753689
[11] Kerstin Bongard-Blanchy, Arianna Rossi, Salvador Rivas, Sophie Doublet, Vincent Koenig, and Gabriele Lenzini. 2021. I am Definitely Manipulated, Even When I am Aware of it. It s Ridiculous!–Dark Patterns from the End-User Perspective. *arXiv preprint arXiv:2104.12653* (2021).
[12] Joseph Bugeja, Andreas Jacobsson, and Paul Davidsson. 2016. On privacy and security challenges in smart connected homes. In *2016 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, 172–175.
[13] Saksham Chitkara, Nishad Gothoskar, Suhas Harish, Jason I Hong, and Yuvraj Agarwal. 2017. Does this app really need my location?: Context-aware privacy management for smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 42.
[14] Isis Chong, Aiping Xiong, and Robert W. Proctor. 2019. Human Factors in the Privacy and Security of the Internet of Things. *Ergonomics in Design* 27, 3 (2019), 5–10. https://doi.org/10.1177/1064804617750321 arXiv:https://doi.org/10.1177/1064804617750321
[15] Juliet Corbin and Anselm Strauss. 2008. Strategies for qualitative data analysis. *Basics of Qualitative Research. Techniques and procedures for developing grounded theory* 3 (2008).
[16] Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. 2006. User interfaces for privacy agents. *ACM Transactions on Computer-Human Interaction (TOCHI)* 13, 2 (2006), 135–178.
[17] Ry Crist. 2019. Amazon and Google are listening to your voice recordings. Here's what we know about that. *c|net* (Jul 2019). https://www.cnet.com/news/best-smart-home-devices-of-2021-that-arent-made-by-amazon-or-google/
[18] F. El-Moussa D. Bastos, M. Shackleton. 2018. Internet of Things: A Survey of Technologies and Security Risks in Smart Home and City Environments. *IET Conference Proceedings* (January 2018), 30 (7 pp.)–30 (7 pp.)(1). https://digital-library.theiet.org/content/conferences/10.1049/cp.2018.0030
[19] Nora A Draper. 2017. From privacy pragmatist to privacy resigned: challenging narratives of rational choice in digital privacy debates. *Policy & Internet* 9, 2 (2017), 232–251.
[20] Janna Lynn Dupree, Richard Devries, Daniel M Berry, and Edward Lank. 2016. Privacy personas: Clustering users via attitudes and behaviors toward security practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 5228–5239.
[21] Janna Lynn Dupree, Richard Devries, Daniel M. Berry, and Edward Lank. 2016. Privacy Personas: Clustering Users via Attitudes and Behaviors toward Security Practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5228–5239. https://doi.org/10.1145/2858036.2858214
[22] Adrienne Porter Felt, Elizabeth Ha, Serge Egelman, Ariel Haney, Erika Chin, and David Wagner. 2012. Android permissions: User attention, comprehension, and behavior. In *Proceedings of the eighth symposium on usable privacy and security*. ACM, 3.
[23] Directorate-General for Communication. 2020 (accessed September, 2020). *EU data protection rules | European Commission.* https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en
[24] Alexandre Fortier and Jacquelyn Burkell. 2018. Display and control in online social spaces: Towards a typology of users. *New Media & Society* 20, 3 (2018), 845–861. https://doi.org/10.1177/1461444816675184 arXiv:https://doi.org/10.1177/1461444816675184
[25] Batya Friedman, Peter H Kahn Jr, Jennifer Hagman, Rachel L Severson, and Brian Gill. 2006. The watcher and the watched: Social judgments about privacy in a public place. *Human-Computer Interaction* 21, 2 (2006), 235–272.
[26] Andrew Gambino, Jinyoung Kim, S Shyam Sundar, Jun Ge, and Mary Beth Rosson. 2016. User disbelief in privacy paradox: Heuristics that determine disclosure. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2837–2843.
[27] Howard Gardner and Katie Davis. 2013. *The app generation: How today's youth navigate identity, intimacy, and imagination in a digital world.* Yale University Press.
[28] State Of California Department Of Justice Office Of The Attorney General. 2020 (accessed September, 2020). *California Consumer Privacy Act (CCPA) | State of California - Department of Justice - Office of the Attorney General.* https://oag.ca.gov/privacy/ccpa
[29] Yuan Gong and Christian Poellabauer. 2018. Protecting voice controlled systems using sound source identification based on acoustic cues. In *2018 27th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 1–9.
[30] Nathaniel S. Good, Jens Grossklags, Deirdre K. Mulligan, and Joseph A. Konstan. 2007. Noticing Notice: A Large-Scale Experiment on the Timing of Software License Agreements. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 607–616. https://doi.org/10.1145/1240624.1240720

[31] Stacey Gray. 2016. Always On: Privacy Implications Of Microphone-Enabled Devices.
[32] Eszter Hargittai and Alice Marwick. 2016. "What can I really do?" Explaining the privacy paradox with online apathy. *International Journal of Communication* 10 (2016), 21.
[33] Yangyang He, Paritosh Bahirat, Bart P Knijnenburg, and Abhilash Menon. 2019. A Data-Driven Approach to Designing for Privacy in Household IoT. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 1 (2019), 1–47.
[34] Karey Helms. 2017. Leaky objects: Implicit information, unintentional communication. In *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*. 182–186.
[35] Christian Pieter Hoffmann, Christoph Lutz, and Giulia Ranzini. 2016. Privacy cynicism: A new approach to the privacy paradox. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 10, 4 (2016).
[36] Jason I Hong, Jennifer D Ng, Scott Lederer, and James A Landay. 2004. Privacy risk models for designing privacy-sensitive ubiquitous computing systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*. 91–100.
[37] Gordon Hull. 2015. Successful failure: what Foucault can teach us about privacy self-management in a world of Facebook and big data. *Ethics and Information Technology* 17, 2 (2015), 89–101.
[38] Luke Hutton and Tristan Henderson. 2017. Beyond the EULA: Improving consent for data mining. In *Transparent Data Mining for Big and Small Data*. Springer, 147–167.
[39] Sasha Ingber. 2018. Amazon Customer Receives 1,700 Audio Files Of A Stranger Who Used Alexa. https://www.npr.org/2018/12/20/678631013/amazon-customer-receives-1-700-audio-files-of-a-stranger-who-used-alexa
[40] Md Tamzeed Islam, Bashima Islam, and Shahriar Nirjon. 2017. Soundsifter: Mitigating overhearing of continuous listening devices. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 29–41.
[41] Spyros Kokolakis. 2017. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & security* 64 (2017), 122–134.
[42] Bert-Jaap Koops, Bryce Clayton Newell, Tjerk Timan, Ivan Skorvanek, Tomislav Chokrevski, and Masa Galic. 2016. A typology of privacy. *U. Pa. J. Int'l L.* 38 (2016), 483.
[43] Sacha Krstulović. 2018. Audio event recognition in the smart home. In *Computational Analysis of Sound Scenes and Events*. Springer, 335–371.
[44] Ponnurangam Kumaraguru and Lorrie Faith Cranor. 2005. *Privacy indexes: a survey of Westin's studies*. Carnegie Mellon University, School of Computer Science, Institute for ....
[45] Josepfphine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, are you listening?: Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 102.
[46] 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*. 501–510.
[47] Ewa Luger, Stuart Moran, and Tom Rodden. 2013. Consent for All: Revealing the Hidden Complexity of Terms and Conditions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2687–2696. https://doi.org/10.1145/2470654.2481371
[48] Sapna Maheshwari. 2018. That Game on Your Phone May be Tracking What You're Watching on TV - The New York Times. (2018).
[49] Nathan Malkin, Joe Deatrick, Allen Tong, Primal Wijesekera, Serge Egelman, and David Wagner. 2019. Privacy Attitudes of Smart Speaker Users. *Proceedings on Privacy Enhancing Technologies* 2019, 4 (2019), 250–271.
[50] Emily McReynolds, Sarah Hubbard, Timothy Lau, Aditya Saraf, Maya Cakmak, and Franziska Roesner. 2017. Toys that listen: A study of parents, children, and internet-connected toys. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 5197–5207.
[51] Christopher Mele. 2016. Bid for Access to Amazon Echo Audio in Murder Case Raises Privacy Concerns. https://www.nytimes.com/2016/12/28/business/amazon-echo-murder-case-arkansas.html
[52] Nicholas Micallef, Mike Just, Lynne Baillie, and Maher Alharby. 2017. Stop annoying me!: an empirical investigation of the usability of app privacy notifications. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*. ACM, 371–375.
[53] Jessica K. Miller, Batya Friedman, Gavin Jancke, and Brian Gill. 2007. Value Tensions in Design: The Value Sensitive Design, Development, and Appropriation of a Corporation's Groupware System. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work* (Sanibel Island, Florida, USA) *(GROUP '07)*. Association for Computing Machinery, New York, NY, USA, 281–290. https://doi.org/10.1145/1316624.1316668
[54] Stuart Moran, Ewa Luger, and Tom Rodden. 2014. An Emerging Tool Kit for Attaining Informed Consent in UbiComp. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (Seattle, Washington) *(UbiComp '14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 635–639. https://doi.org/10.1145/2638728.2641677
[55] Anthony Morton and M. Angela Sasse. 2014. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*. 102–111. https://doi.org/10.1109/PST.2014.6890929

[56] Anthony Morton and M Angela Sasse. 2014. Desperately seeking assurances: Segmenting users by their information-seeking preferences. In *2014 Twelfth Annual International Conference on Privacy, Security and Trust*. IEEE, 102–111.

[57] Najmeh Mousavi Nejad, Simon Scerri, and Sören Auer. 2017. Semantic Similarity Based Clustering of License Excerpts for Improved End-User Interpretation. In *Proceedings of the 13th International Conference on Semantic Systems* (Amsterdam, Netherlands) *(Semantics2017)*. Association for Computing Machinery, New York, NY, USA, 144–151. https://doi.org/10.1145/3132218.3132224

[58] Najmeh Mousavi Nejad, Simon Scerri, Sören Auer, and Elisa M. Sibarani. 2016. EULAide: Interpretation of End-User License Agreements Using Ontology-Based Information Extraction. In *Proceedings of the 12th International Conference on Semantic Systems* (Leipzig, Germany) *(SEMANTiCS 2016)*. Association for Computing Machinery, New York, NY, USA, 73–80. https://doi.org/10.1145/2993318.2993324

[59] David H. Nguyen, Gabriela Marcu, Gillian R. Hayes, Khai N. Truong, James Scott, Marc Langheinrich, and Christof Roduner. 2009. Encountering SenseCam: Personal Recording Technologies in Everyday Life. In *Proceedings of the 11th International Conference on Ubiquitous Computing* (Orlando, Florida, USA) *(UbiComp '09)*. ACM, New York, NY, USA, 165–174. https://doi.org/10.1145/1620545.1620571

[60] Sarah Perez. 2019. Google is investigating the source of voice data leak, plans to update its privacy policies. (Jul 2019). https://techcrunch.com/2019/07/11/google-is-investigating-the-source-of-voice-data-leak-plans-to-update-its-privacy-policies/

[61] Anne Pfeifle. 2018. Alexa, What Should We Do about Privacy: Protecting Privacy for Users of Voice-Activated Devices Comments. *Washington Law Review* 93 (2018), 421.

[62] James Pierce. 2019. Smart home security cameras and shifting lines of creepiness: A design-led inquiry. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[63] Lee Rainie and Kathryn Zickuhr. 2015. Americans' views on mobile etiquette. *Pew Research Center* 26 (2015), 948–958.

[64] Tom Rodden and Steve Benford. 2003. The evolution of buildings and implications for the design of ubiquitous domestic environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 9–16.

[65] Manuel Polst Rudolph, Denis Feth, et al. 2019. Usable Specification of Security and Privacy Demands: Matching User Types to Specification Paradigms. *Mensch und Computer 2019-Workshopband* (2019).

[66] Ali Salman. 2020. iOS 14 Will Notify You With Green and Orange Indicators When Apps Use Microphone or Camera. https://wccftech.com/ios-14-will-notify-you-with-green-and-orange-indicators-when-apps-use-microphone-or-camera/

[67] Oliver Schneider and Alex Garnett. 2011. ConsentCanvas: Automatic Texturing for Improved Readability in End-User License Agreements. In *Proceedings of the ACL 2011 Student Session* (Portland, Oregon) *(HLT-SS '11)*. Association for Computational Linguistics, USA, 41–45.

[68] Eva-Maria Schomakers, Chantal Lidynia, Luisa Vervier, and Martina Ziefle. 2018. Of Guardians, Cynics, and Pragmatists-A Typology of Privacy Concerns and Behavior. In *IoTBDS*. 153–163.

[69] Eva-Maria Schomakers, Chantal Lidynia, and Martina Ziefle. 2019. A Typology of Online Privacy Personalities. *Journal of Grid Computing* 17, 4 (2019), 727–747.

[70] Kim Bartel Sheehan. 2002. Toward a typology of Internet users and online privacy concerns. *The Information Society* 18, 1 (2002), 21–32.

[71] Fuming Shih, Ilaria Liccardi, and Daniel Weitzner. 2015. Privacy tipping points in smartphones privacy preferences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 807–816.

[72] Jay Stanley. 2017. The Privacy Threat From Always-On Microphones Like the Amazon Echo. https://www.aclu.org/blog/privacy-technology/privacy-threat-always-microphones-amazon-echo

[73] Linda Sui. 2016. Strategy Analytics: Android captures record 88 percent share of global smartphone shipments in Q3 2016. *Strateg. Anal. Res. Experts Anal* 28 (2016), 28–35.

[74] Madiha Tabassum, Tomasz Kosiński, Alisa Frik, Nathan Malkin, Primal Wijesekera, Serge Egelman, and Heather Richter Lipford. 2019. Investigating Users' Preferences and Expectations for Always-Listening Voice Assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 4, Article 153 (Dec. 2019), 23 pages. https://doi.org/10.1145/3369807

[75] Joshua Tan, Khanh Nguyen, Michael Theodorides, Heidi Negrón-Arroyo, Christopher Thompson, Serge Egelman, and David Wagner. 2014. The effect of developer-specified explanations for permission requests on smartphone user behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 91–100.

[76] Joseph Turow, Michael Hennessy, and Nora Draper. 2015. The tradeoff fallacy: How marketers are misrepresenting American consumers and opening them up to exploitation. *Available at SSRN 2820060* (2015).

[77] T. Franklin Waddell, Joshua R. Auriemma, and S. Shyam Sundar. 2016. Make It Simple, or Force Users to Read? Paraphrased Design Improves Comprehension of End User License Agreements. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5252–5256. https://doi.org/10.1145/2858036.2858149

[78] Paul Wagenseil. 2021. Zoom security issues: Here's everything that's gone wrong (so far). *Tom's Guide* (Jan 2021). https://www.tomsguide.com/news/zoom-security-privacy-woes

[79] Ari Ezra Waldman. 2019. There is No Privacy Paradox: How Cognitive Biases and Design Dark Patterns Affect Online Disclosure. *Current Opinion in Psychology* (2019).

[80] Xuetao Wei, Lorenzo Gomez, Iulian Neamtiu, and Michalis Faloutsos. 2012. Permission evolution in the android ecosystem. In *Proceedings of the 28th Annual Computer Security Applications Conference*. ACM, 31–40.

[81] Jacob O. Wobbrock, Shaun K. Kane, Krzysztof Z. Gajos, Susumu Harada, and Jon Froehlich. 2011. Ability-Based Design: Concept, Principles and Examples. *ACM Trans. Access. Comput.* 3, 3, Article 9 (April 2011), 27 pages. https://doi.org/10.1145/1952383.1952384

[82] Sam Wolfson. 2018. Amazon's Alexa recorded private conversation and sent it to random contact. https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation

[83] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. 2019. Defending My Castle: A Co-Design Study of Privacy Mechanisms for Smart Homes. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, NY, USA, Article 198, 12 pages. https://doi.org/10.1145/3290605.3300428

[84] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 200 (Nov. 2018), 20 pages. https://doi.org/10.1145/3274469

[85] Shoshana Zuboff. 2018. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1st ed.).