# Context-Sensitive Inquiry Through the Lens of Stakeholders: Privacy and Security Problems with Emerging Technologies

**Andrea Gallardo**

CMU-S3D-25-117

September 2025

Software and Societal Systems Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

**Thesis Committee:**
Lujo Bauer, ECE/S3D, Co-chair
Lorrie Cranor, EPP/S3D, Co-chair
Emma Strubell, LTI
Renee Shelby (Google, Inc.)
Robert Erbes (Idaho National Laboratory)

*Submitted in partial fulfillment of the requirements*
*for the degree of Doctor of Philosophy in Societal Computing.*



This research was sponsored or supported in part by the GEM Fellowship, Idaho National Laboratory, Meta, Google, Highmark, the Defense Technology Security Administration, the Office of Naval Research, the NSF Center for Distributed Confidential Computing (grant CNS-2207216), and S3D departmental funding from the Software and Societal Systems Department in the School of Computer Science at Carnegie Mellon University.

The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies or positions, either expressed or implied, of any sponsoring institution.

**Keywords:** Privacy, Security, Human Computer Interaction, Emerging Technologies

# Abstract

Human-centered research on emerging technological applications can inform the development of secure, privacy-preserving, products and procedures that align with social and legal standards. Directly engaging with stakeholders situated in particular contexts can provide detailed perspectives that reveal limitations, problems, or unmet needs in security approaches, data collection, data use, and accuracy of technologies.

This thesis demonstrates how direct engagement with stakeholders illuminates crucial problems and perspectives, consisting of four research studies consulting potential users or data subjects of technologies with broad social implications, focusing on four emerging technological applications: 1) AR glasses with advanced sensor capabilities, 2) AI analysis of voice data for decision-making in employment and education, 3) the convergence of information technology (IT) and operational technology (OT) in energy grid infrastructure, and 4) patient-facing medical translation tools. We first consider current AR users' context-sensitive privacy attitudes, preferences, and concerns regarding how future hypothetical AR glasses could collect and use data. Second, we report on how speakers of four US English dialects perceived potential benefits and harms of AI-enabled voice analysis in high-stakes employment and educational contexts. Third, we compare approaches to vulnerability impact assessment among experts in critical infrastructure and computer security, identifying notable differences in the self-reported approaches to assessing risk in energy systems between these two groups. Finally, we present attitudes, preferences, and concerns of Mandarin- and Spanish-speaking individuals with limited English proficiency (LEP) towards existing and emerging translation services and technologies in medical contexts.

We contribute an interdisciplinary approach to eliciting and analyzing firsthand stakeholder insights about privacy and security problems with emerging technologies. We use human-computer interaction methods to obtain rich, qualitative data and apply thematic coding, discourse analysis, and sociocultural anthropological considerations to contextualize participants' responses. Our approach elucidates context-sensitive and socially situated perceptions and attitudes regarding the privacy and security of various technologies, highlighting social dimensions often overlooked in usable privacy and security literature and revealing broader implications for policy and technology design.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Consulting stakeholders of emerging technologies can shed light on potential privacy and security problems posed by these technologies, as well as societal and policy implications. Taking proactive approaches to solicit context-specific insights and socially situated concerns of end-users and data subjects can expose how problems such as privacy violations, computer vulnerability exploits, or algorithmic errors can cause harm with far-reaching social impacts, including, as conveyed in this thesis, threatening legal rights, social norms, critical infrastructure security, and patient understanding of medical care. Human-centered investigations of existing and potential problems is thus essential for developing secure, privacy-protecting, and socially acceptable technologies.

We contribute a set of human-centered qualitative research studies with interdisciplinary data collection and analysis of privacy and security problems with emerging technologies. Using HCI and usable privacy and security (UPS) methods, we employ semi-structured interviews and a survey to engage stakeholders and elicit attitudes, preferences, concerns, and expectations informed by their lived experiences. In our final three studies, we apply sociocultural anthropological considerations throughout the research process, which enables us, as well as participants, to make connections to emerging technologies or practices that might go unexplored in a broader study. As we are not conducting ethnographies of specific populations in situ, we enable such cultural analysis by intentionally choosing populations, use cases, or locations that ground our investigations in particular social, cultural, and linguistic contexts.

To interpret our data, we employ thematic coding and discourse analysis to categorize participants' responses, extracting themes that align with our research questions about privacy, security, and algorithmic harms. In addition to using typical HCI and UPS qualitative analysis methods, we apply sociocultural anthropology considerations, keeping in mind factors such as location, language, social roles and relationships, and behavioral and institutional norms. We also consider participants' self-reported lived experiences, expertise, characteristics, and other sociocultural factors when analyzing the data. Additionally, we acknowledge our own practice of interpretation and meaning-making and the academic and epistemological influences on our thematic coding processes. Each participant's response is an act of meaning-making, and the researcher who receives it and analyzes it, in turn, also engages in an act of meaning-making.

The four studies described in this thesis illustrate this interdisciplinary approach. We focus on four emerging applications of technologies: hypothetical future AR glasses, AI analysis of voice

data in education and employment decision-making use cases, IT/OT convergence in energy critical infrastructure, and machine translation in medical settings. Our results, based on qualitative analyses of participant responses, demonstrate that context-specific insights can reveal important socially relevant considerations for the design and deployment of emerging technologies. These insights include privacy, security, fairness, and accuracy problems that stakeholders anticipate, their attitudes, preferences, and concerns, and connections they make to experiences, history, and cultures relevant to them.

In our final study, we applied this interdisciplinary approach to interview speakers of languages other than English (LOE), Mandarin and Spanish, and elicit their attitudes and preferences for using existing and emerging translation technologies and services in medical contexts. We ground our analysis in their self-reported experiences with translation and interpretation services and technologies in the local area and in broader social and policy contexts of medical language access in Pennsylvania.

Our work documents expectations for technical systems with broad and varied social implications. Elucidating context-specific stakeholder insights into emerging technologies is especially pertinent use cases for which regulations and social norms are still developing. As part of each study, we make recommendations for policy makers and technology professionals, such as enacting privacy laws and developing auditing practices for AI systems.

Overall, this thesis provides a deeper understanding of conducting interdisciplinary human factors research in cybersecurity, digital privacy, and ethics for emerging technological contexts by interviewing and surveying stakeholders about context-specific use cases and conducting qualitative analysis of their self-reported privacy and security attitudes, concerns, experiences, and approaches. We provide valuable insights into how such individual and context-specific data can shed light on policy implications and the potential societal impact of emerging technological applications. Based on these insights, we make recommendations for policymakers, industry professionals, and technology designers.

## 1.1 Thesis Statement

We contribute a set of human-centered qualitative research studies with interdisciplinary data collection and analysis of firsthand stakeholder insights about privacy and security problems with emerging technologies. We use human-computer interaction methods to obtain rich, qualitative data and apply thematic coding, discourse analysis, and sociocultural anthropological considerations to contextualize participants' responses. Our approach elucidates context-sensitive and socially situated perceptions and attitudes regarding the privacy and security of various technologies, highlighting social dimensions often overlooked in usable privacy and security literature and revealing broader implications for policy and technology design.

## 1.2 Roadmap

In the following chapters, we present four qualitative studies that exemplify our human-centered and interdisciplinary research approach to studying stakeholder perceptions and concerns about

emerging technologies and contexts. Specifically, we first present an interview study on the potential privacy concerns of current AR users regarding the collection and use of AR glasses data. We then present an interview study on an emerging technological security problem: the integration of digital and information technology (IT) components and software into energy operational technology (OT) used to operate the energy grid. We then consider a survey study considering US southerners' perceived benefits and harms of AI analysis of voice data in high-stakes employment and education contexts. completing Finally, we present a study on the use of AI-augmented machine translation in medical contexts.

In the overview section of each study, we draw connections between the works, emphasizing the context-dependent nature of stakeholder considerations regarding emerging technological applications and their potential social implications. Below, I provide a summary of the studies included in the thesis.

### 1.2.1 Speculative Augmented Reality Glasses Data Collection

Our PoPETS '23 study examined the attitudes and perceptions of a specific population: current users of augmented reality (AR). Participants' self-reported comfort levels and concerns about various current and potential AR glasses data collection features, such as video recording or facial recognition, highlight the context-dependent nature of privacy concerns and the need for customizable privacy solutions that meet users' diverse needs.

### 1.2.2 Emerging IT/OT Convergence in Critical Infrastructure

Our CHI '24 study examined and compared the computer vulnerability impact assessment approaches of two kinds of experts working in the energy sector: computer security researchers and energy operations professionals. Our findings underscore the imperative to establish cross-domain knowledge and cultivate collaboration in order to build a better understanding of cybersecurity risk in power systems. Our interdisciplinary approach helps shed light on organizational cultural issues and epistemological challenges in securing critical infrastructure.

### 1.2.3 AI Analysis of Voice Data in High Stakes Contexts

This survey study (in submission) explores the social and ethical implications of AI analysis of voice data in employment and education contexts, considering potential social and legal consequences of using voice technologies to evaluate data subjects in hiring, job performance evaluation, college admissions, and exam grading contexts. We asked participants to consider sociolinguistic differences as speakers of different American English dialects from the southern US. While over half of the participants found the use of AI for decision-making acceptable, a vast majority emphasized the potential for bias and discrimination.

### 1.2.4 Translation-Mediated Disclosure

This interview study explores language access experiences and needs of Mandarin- and Spanish-speaking individuals who are primarily speakers of languages other than English (LOE) with

limited English proficiency (LEP) in medical settings. We asked participants to consider existing services and tools as well as a hypothetical advanced AI health assistant that could assist them with translation or interpretation in medical situations, and we report their experiences, perceptions and preferences regarding the use of such services and technologies.

# Chapter 2

# Speculative Privacy Concerns about AR Glasses Data Collection

This chapter was adapted from my published paper:

> Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie Cranor. 2023. "Speculative Privacy Concerns about AR Glasses Data Collection." *Proceedings on Privacy Enhancing Technologies. (PoPETS '23)*
> `https://petsymposium.org/popets/2023/popets-2023-0117.php`

## 2.1 Overview

In this chapter, we present a study on the attitudes of current AR users towards the potential data collection and data use capabilities of advanced augmented reality (AR) glasses. We report results from 21 semi-structured interviews with current AR users regarding hypothetical data collection and use cases for speculative future AR glasses, in which participants discussed whether they would be comfortable or uncomfortable with the collection of 15 different types of data and five use cases, often emphasizing the context-dependent nature of their privacy attitudes, concerns, and preferences.

This work relates to the thesis through its detailed qualitative analysis of participants' context-sensitive considerations, which reveals a potential disconnect between advanced sensor and analysis capabilities and existing social norms and personal preferences. By prompting participants with specific data collection features and use cases, we were able to elicit unique and detailed reactions that touched on potential societal implications. We utilized existing frameworks of Helen Nissenbaum's philosophy of contextual integrity and Daniel Solove's privacy harm taxonomy to connect our thematic coding and discourse analysis of participants' responses to broader concepts of social norms and privacy violations.

Some of the rich examples of potentially problematic applications are directly relevant to the following work in the thesis on AI analysis of voice data, including concerns about discrimination (12 participants suggested that discrimination could result from normative biases built into AR

glasses features) and responses relating to algorithmic evaluation of speech and gestures, social and conversational feedback, and emotion or mood detection and feedback, which raised the possibility of inaccurate or pseudoscientific evaluations. For example, P20 imagined a social feedback feature negatively assessing their non-standard American English accent from "rural Appalachia": "If the glasses are telling me that my accent is poor and that I need to retrain myself how to speak, that, I would have a bit of a bigger issue with." Similar to respondents in our study on AI-enabled analysis of voice data who emphasized that voice characteristics such as dialect do not determine how skilled an employee or student is, P20 said that "accent does not determine how intelligent or capable another person is." Additionally, P7 expressed concern that facial expression and gesture analysis may not work well for people who do not express themselves in normative ways, such as some autistic and neurodiverse people, and the potential "misuse" by employers of such algorithmic analysis. Other participants also suggested it could be harmful for employers to collect and use gestural, social feedback, reaction time, or emotion and mood data, with P12 anticipating uneven benefits: "An employer, looking at the reaction time of all their employees and saying this person's really slow, they're getting fired, right, it seems to benefit external third parties more than the [data subject]."

These results, which convey the importance of considering context and social groups, provided a foundation to conduct contextualized and population-specific studies on privacy and security problems with emerging technologies through the lens of end users and data subjects.

## 2.2 Abstract

As technology companies develop mass market augmented reality (AR) glasses that are increasingly sensor-laden and affordable, uses of such devices pose potential privacy and security problems. Though prior work has broadly addressed some of these problems, our work specifically addresses the potential data collection of 15 data types by AR glasses and five potential data uses. Via semi-structured interviews, we explored the attitudes and concerns of 21 current AR technology users regarding potential data collection and data use by hypothetical consumer-grade AR glasses. Participants expressed diverse concerns and suggested potential limits to AR data collection and use, evoking privacy concepts and informational norms. We discuss how participants' attitudes and reservations about data collection and use, like definitions of privacy, are varying and context-dependent, and make recommendations for designers and policy makers, including customizable and multidimensional privacy solutions.

## 2.3 Introduction

Augmented reality (AR) glasses pose privacy concerns, with their potential to collect sensitive user data—such as biometrics (heart rate, eye tracking, voiceprint) and bystander face images—and to combine such data to infer even more sensitive user or bystander characteristics, such as health status. The failure of Google Glass, Google's AR glasses, a decade ago is sometimes attributed to a lack of consideration for societal norms and privacy expectations [116, 160]. As information norms for consumer-grade AR devices are still being established, a better under-

standing of context-relevant privacy risks and concerns can help inform the design of these technologies and help ensure that they respect privacy.

We explore potential privacy and security concerns regarding AR glasses data collection through 21 semi-structured interviews with current users of AR technologies available to general consumers. We told participants to imagine they had a more advanced (hypothetical) pair of AR glasses and asked them how comfortable they would be with the collection of 15 types of data and five data use cases. Our research questions were as follows: (RQ1) What are participants' attitudes and concerns regarding data collection and use by future AR glasses? (RQ2) Which data actors and data subjects are participants concerned about? (RQ3) What privacy principles do they expect or hope for?

We used qualitative coding methods to analyze the interview data, using a priori attitude labels and also iteratively developing emergent codes. Most participants were uncomfortable or had mixed feelings about certain data types, such as face images, brain waves, and voiceprints. They also had concerns regarding potential data actors, such as data collectors or receivers (e.g., employers, advertisers, and doctors) and data subjects, such as bystanders or their children. Participants also presented potential context-relevant privacy harms as well as potential harms to people in vulnerable situations or groups. Privacy principles desired by participants included control over data collection and use, providing notice, requiring consent, collecting or using only necessary information, and protecting sensitive information.

Our exploratory study provides insight into current AR technology users' privacy concerns regarding data collection by future AR glasses and their desired privacy principles. From these insights we offer recommendations for AR professionals and lawmakers.

## 2.4 Background and Related Work

Below we discuss AR glasses technology, including current and future data collection capabilities, as well as prior work on privacy and security concerns about AR, mixed reality (MR), and virtual reality (VR) glasses, and related technologies. We outline some privacy concepts or frameworks, which we use to analyze participants' responses in our study, and prior work on privacy risks for people in vulnerable situations or groups in VR and online contexts.

### 2.4.1 Background on AR Glasses

Augmented reality (AR) is a technology that augments a user's visual and audio perception of reality with interactive virtual overlays. AR glasses are wearable head mounted display devices that provide such functionalities and are different from other devices that use sensors, such as mobile phones or IoT devices, in their distinctive combination of form factor, functionalities, and sensors. They are sometimes also considered mixed reality (MR) glasses, when they provide the ability to interact with both virtual and physical objects, in contrast to VR glasses, which provide fully virtual interactions.

### 2.4.1 (a) Information Flows for AR/MR/VR Glasses and IoT

We discuss types of data collected and used by AR and MR glasses, as well as by wearable IoT technology, such as smart watches, whose sensors and features may be integrated into AR devices in the future. These helped determine the data types and data uses we asked our participants to consider. Data collected by these devices' sensors often contain personal or sensitive information, or could be combined with other data types to reveal such information. We also anticipate AR glasses integrating mobile device features like notifications, face filters, and location mapping, as some mobile AR apps or features utilize cameras, e.g., to overlay digital graphics on camera images, allowing people to create or adopt face filters, or use location mapping, e.g, for mobile AR gameplay in Pokémon GO.

**AR Glasses Sensors and Features** The number of sensors in AR devices has increased, but as prior work notes, not all components are made known by companies and sensor presence sometimes has to be "inferred from device functionality" or "direct observation by taking it apart or consulting online resources detailing such observation" [247]. Recent mass market AR glasses are equipped with basic sensors, such as microphones, speakers, or cameras, to collect audio and video or image data and to detect surroundings, but have limited (or sometimes no) embedded displays, with overlain images sometimes serving primarily to display video or browser content rather than content that interacts with one's environment [146, 353], Realistic AR integration of overlain objects and visual perception of physical surroundings has not yet been achieved.

Enterprise level AR glasses tend to have additional sensors, cameras, and tools to track a user's eye and body movement, map surroundings (such as indoor spaces), and create 3D models. The HoloLens2 sensors include visible light and infrared cameras, a depth sensor, an inertial measurement unit (IMU), and a camera capable of recording video. It also collects biometric data such as eye tracking and body movement data [28, 276]. In addition to similar sensors, Varjo's XR-3 headset uses light detection and ranging (LiDAR) for depth-sensing [391]. The way that data is stored also varies between devices, e.g., Hololens2 eye-tracking data is stored as eye gaze vectors [275], while Varjo's eye tracking data consists of "foveated rendering" and raw video recordings of eye movement [420]. Apple's yet-to-be-released Vision Pro glasses use 12 cameras, five sensors, and six microphones, including "high-speed cameras and a ring of LEDs" for eye tracking [23].

**IoT sensors and features** We anticipate AR glasses integrating features of wearable IoT devices like fitness trackers and smart watches, which can collect biometric data such as heart rate, body temperature, and movement data, and are used for health monitoring [394]. We also imagined AR glasses containing sensors or features not yet generally marketed to consumers, such as brain wave data collected by EEG sensors (currently integrated into MR glasses as part of Varjo's Galea Beta Program) [139, 140, 200, 392, 418], facial recognition [194, 216], the capture and use of voiceprints [124, 228, 401, 449] and reaction times [360], and feedback about facial expressions [259, 283, 367], mood [134], or social interactions [57].

#### 2.4.1 (b) Privacy Policies for Current AR/VR Glasses

Current privacy policies vary in the amount of detail they provide about AR or VR glasses data collection and contain little to no detail about inferences made using headset data. For example, Google's, HTC Vive's, and Varjo's devices' Terms of Service agreements direct readers to their companies' general privacy policies, which do not mention how device data is used [75, 76, 133, 135, 419, 421]. Meta provides various privacy notices, some specifically addressing data about movement and recording in virtual worlds [266, 267, 268, 269]. In immersive VR settings, the Oculus collects audio recordings "through a rolling buffer processed locally on-device" that can be stored on Meta's servers if a report is submitted to them to report abuse or harmful conduct [268]. Apple provides a general privacy policy as well as "product-specific" policies for features that exist across different Apple devices [21, 22]. Apple Vision Pro promotional materials also state sharing limitations for information about a user's iris, eye tracking information, and data from the camera and other sensors [23].

Yet, as has been observed in prior work [5], no full accounting of data collection practices is provided in these legal notices. Meta and HTC also inform users that separate privacy conditions apply for services or products provided by third parties [75, 76, 267]. Magic Leap and Snap, unlike Google, HTC Vive, or Varjo, provide more detailed privacy policies for their glasses, covering what information is collected, how it is used, and with whom it is shared [172, 174]. While their devices have different types of sensors, e.g., Magic Leap 2 can track users' eyes [173], whereas Snap Spectacles 3 cannot [175], both privacy policies clearly state the data collected from different sensing modalities and, like Meta and Microsoft, mention general uses for collected data: personalizing content, improving user experience, and product development. However, whether and what inferences are made about users with the collected biometric data (e.g., eye tracking, audio) is not stated and remains vague.

### 2.4.2 Privacy and Security Concerns

In this section, we discuss prior work that addresses user concerns about AR glasses and related technologies. Prior work has anticipated potential privacy and security harms arising from AR glasses [47, 224, 357] and AR applications such as mobile AR games [334]. De Guzman et al. outline potential mixed reality privacy and security risks and provide a survey of prior work addressing them [91]. Harborth et al. showed a gap between users' understanding of threat models and actual privacy and security risks of certain mobile AR permissions, such as accelerometer data, which tracks movement [144]. Prior work has also considered potential privacy and security risks of data inference, which users may be unaware of. Cong et al. designed an eavesdropping attack using zero-permission motion sensors in AR/VR glasses to infer speaker gender, identity, and speech content from live human speech [374]. Bye et al. raised concerns that biometric data could be used to infer an MR glasses user's membership in marginalized groups, noting that gaze data can reveal users' sexual preferences [47, 342]. Other work confirms that eye tracking can reveal gender, sexual preference, age, race, affect, emotional state, health, and task focus [212, 236, 432].

Yet, only a few studies have considered current or potential AR glasses users' privacy and security concerns [93, 144, 206, 207, 224, 351]. These explored participants' general feelings

of comfort, acceptability, or concern about unspecified general AR or MR glasses usage, but our study is the first to focus on current AR users' concerns about AR glasses collecting and using specific data types.

#### 2.4.2 (a) AR technology.

In prior work surveying or interviewing end users or potential users of AR technology, participants' privacy and security concerns include AR glasses capturing private information, collecting data about their physical surroundings, biometrics and private activities, bystanders' privacy and security, risks posed by overlain content, and security compromise of AR devices or apps [93, 224, 306]. Koelle et al. conducted two studies gauging acceptability of the general use of "data glasses" and found that social context, such as whether other people were present, was a factor in how acceptable participants found the glasses [206, 207]. Denning et al. found that bystanders in an AR context were concerned about being identified, citing reasons like bodily harm as imagined negative consequences [93]. Rixen et al. addressed potential users' comfort with AR technology displaying personal information (rather than with data collection or use, as in our study), and found that factors such as people present ("intimacy") and whether information was self-disclosed or not influenced participants' comfort level [351]. A survey conducted by O'Hagan et al. showed that participants' privacy wishes and preferences, as hypothetical bystanders to AR glasses' sensing and data collection, varied by feature and was also context dependent, with participants providing examples of situations they would find problematic [306].

#### 2.4.2 (b) Social VR and Online Social Contexts

Prior work has considered privacy and security concerns in the contexts of social VR (VR communities that allow live interaction between users, such as VR Chat, AltspaceVR, or Horizon Worlds) and online communities in which the use of pseudonyms and avatars can protect anonymity and allow users to engage in selective self-disclosure. Studies on self-disclosure in social VR show that while some users expressed feeling more comfortable sharing personal information about emotions, sexuality, lifestyle, and personal goals when using anonymous avatars than in person, many had concerns about disclosing personal information about gender, cultural or linguistic background, or disability status through their voices, avatars' features, or behavioral patterns [249, 399, 450]. On social media, Li et al. found significant differences in female and male privacy disclosures [230], and Pyle et al. shed light on how a user's LGBTQ identity might lead them to not disclose potentially stigmatizing information such as pregnancy loss [328]. Prior work has also noted the potential for VR to both reinforce and mitigate social prejudice [403].

However, users, especially those from non-dominant groups, face privacy and security risks that may discourage participation [122]. Prior work has shown that online harassment or suppression of marginalized community members by dominant group members has occurred in online gaming communities such as Second Life and Massively Multiplayer Online Role-Playing Games [30, 80, 119, 263, 402], as well as on social media [292, 407]. VR and AR technologies allow for more immersive interactions, e.g., the ability to simulate physical interaction and intimacy, which has also led to simulations of sexual harrassment and assault [249, 374]. Some scholars have also suggested that norms based on a majority or dominant group's practices may

result in exclusion and increased marginalization of people who cannot or do not conform to these existing norms [262]. In a study on VR in the context of disability, Gerling et al. argue that the design of VR technology assumes a normative "corporeal standard" and consequently excludes disabled people by failing to adequately accommodate them [128]. In our study, participants identified similar possibilities for exclusion, harm or discrimination.

#### 2.4.2 (c) IoT and Wearable Cameras and Recording

Prior work in the context of IoT has explored concerns related to wearable cameras, such as lifelogging cameras [162, 163, 323], as well as photo and video recording in public [326], and private [163] spaces. People have also expressed concerns about bystander privacy being violated by cameras [118, 162, 326]. Considering that the frame rates of traditional AR head-mounted displays are likely higher than some of these lifelogging cameras [162], more fine-grained data collection and possible combinations of data for inference may pose privacy issues.

### 2.4.3 Attitudes and Comfort Levels

Prior work capturing user attitudes towards privacy has used comfort as a proxy for how participants feel about a given topic [152, 193, 245], including AR [145, 220, 351]. Some of this work [145] focused on developing scales that measure how much benefits of using AR outweigh privacy concerns or vice versa. In contrast, our study uses qualitative methods to elicit more and richer detail than such scales could capture about potential privacy concerns, through the lens of participants' feelings of comfort or acceptability and discomfort or non-acceptability of AR glasses data collection or use.

### 2.4.4 Privacy Concepts

Here, we note various concepts of privacy, which we will apply to participants' responses about data collection and use, in Section 2.9. We also note some shortcomings of these concepts.

Some common conceptions of privacy include privacy as seclusion (freedom from intrusion), control over personal information (ability to control information flows), and confidentiality (preventing unwanted disclosure or exposure) [14, 153, 363]. Scholars have suggested that there are private spheres and public spheres, distinct spaces or contexts delineating privacy boundaries [288]. Yet, there is no single definition of privacy, and critics of such privacy concepts point out that they are not able to capture the diversity and complexity of privacy across contexts and societies [14, 153, 290, 383]. Hartzog writes, "When lawmakers and judges accept privacy as a concept that contains multitudes, each of these different notions can explicitly be brought to bear on the real needs of people, groups, and institutions rather than deploying an ill-fitting theory in diverse contexts" [153]. For example, privacy as control has inspired frameworks such as notice and choice, but such "privacy self-management" approaches make individuals responsible for protecting themselves even in the face of overwhelming amounts of privacy policies or insurmountable imbalances of power [153, 384].

Nissenbaum and Solove have proposed frameworks for considering the practical implications of privacy violations. Solove's work considers privacy problems and harms using a taxonomy

with four categories of problems, i.e., "four basic groups of harmful activities": (1) information collection, (2) information processing, (3) information dissemination, and (4) invasion. Each of these groups consists of different related subgroups of harmful activities. In this study, we apply a subset of Solove's taxonomy of privacy harms to the privacy concerns voiced by participants [383]. These include the subgroups of aggregation, appropriation, breach of confidentiality, disclosure, distortion, exclusion, exposure, identification, intrusion, secondary use, and surveillance. Prior work has mapped Solove's taxonomy to participant concerns in interview studies, focusing on connecting them to the taxonomy's four categories of problems [12, 61, 335]. While our work focuses on concerns about data collection, participants also brought up concerns spanning these four categories.

Nissenbaum's norms-based approach to privacy, known as contextual integrity, is concerned with "context-relative informational norms," which are social norms "specifically concerned with the flow of personal information." Contextual integrity requires that "practices affecting information flows be assessed in terms of their compliance with context-relevant information norms," which can be evaluated in terms of "values, purposes, and goals." She suggests that informational norms that regulate social behavior (e.g., acceptable actions or practices) in contexts of social life should also exist in digital contexts, and that these norms should be discussed and established by all stakeholders, rather than having arbitrary rules stated by companies [288, 289, 290]. Scholars have used Nissenbaum's Contextual Integrity framework to investigate whether privacy attitudes or expectations align with existing data flows. Apthorpe et al. used surveys to study how participants' attitudes aligned with data regulations [24]. Zhang et al.'s qualitative analysis of survey responses revealed that nuanced ethical and social values informed participants' "normative assessment of the perceived appropriateness of" a given technology [451]. Participants in our study often raised context-relative privacy considerations and emphasized contextual factors, such as physical location, social context, or data actors involved in the information flows. We discuss potential connections to contextual integrity in Section 2.9.2.

## 2.5 Methods

Our study consisted of 21 semi-structured interviews with current AR technology users and was approved by our institution's IRB. We describe recruitment, interview, and data analysis methods below.

### 2.5.1 Recruitment

We recruited participants with varied prior AR experiences through posts to Reddit forums and an email listserv related to new media (e.g., AR), seeking permission from forum moderators to post our recruitment text (Appendix A.1). Specifically, we recruited from forums for general AR or Hololens users (r/hololens, r/augmentedreality) and forums for fans of mobile AR games Ingress, Pokémon GO, and Harry Potter Wizards Unite (r/ingress, r/pokemongo, r/hpwu). The breakdown of recruitment sources by participant is included in Appendix A.4. We paid participants $20 for 60-minute interviews and $30 for 90-minute interviews.

| **Data Types and Uses of AR Glasses** |
|---|
| **Data Types** |
| Audio |
| Video or Image |
| Location |
| Indoor Spaces |
| Virtual Spaces |
| Heart Rate |
| Body Temperature |
| Brain Waves |
| Movement |
| Eye Tracking |
| Face Images |
| Expressions |
| User Voiceprint |
| Bystander Voiceprint |
| Reaction Times |
| **Data Uses** |
| **Notifications:** reminders or notices alerting the user |
| **Health Monitoring:** health feedback based on biometric data |
| **Social Feedback:** based on interaction data, gesture, or voice |
| **Face Filter:** overlain images used as avatars or accessories |
| **Mood/Emotions:** predicted emotional state (e.g., based on tone) |

Table 2.1: A list of 15 data types collected by hypothetical AR glasses and five data uses.

We asked all potential participants to fill out a screening survey to ensure they spoke and understood English, were located in the U.S. were at least 18 years old, were able to install and run Zoom for the interview, and had used at least one AR app or device recently. We purposefully invited participants of various gender, racial, and ethnic identities from among those who filled out the screening survey to participate in the study. We anticipated that users with prior experience using current AR technology would be more knowledgeable about what data can be collected and more aware of possible privacy and security risks than other people, and might also have insights about using AR devices that travel with them and collect data in varied social contexts (e.g., work, education).

### 2.5.2 Interviews

Using an interview format (21 semi-structured interviews) instead of a survey allowed us to elicit participant responses that better captured nuances, conditionals, and mixed or conflicted opinions. We piloted the study with five participants and then revised the protocol substantially to focus our questions more on data collection. We piloted the revised study with two additional participants,

who completed it in under 60 minutes.

Each interview took place over Zoom and was recorded and automatically transcribed after obtaining consent to record. Only one participant chose to leave their video on during the recording. The first 11 interviews took up to 60 minutes, and the last 10 interviews took between 60–90 minutes. We extended the interview length, as we were often unable to complete it in 60 minutes.

The interviews started with questions regarding background information on prior AR technology use and participants' understanding of data collection practices of the AR technology they most often use. The main part of the study consisted of questions regarding participants' attitudes toward data collection and use by We asked participants whether they would be comfortable or uncomfortable with the collection of 15 specific data types and five specific data uses (listed in Table 2.1). For the data use of health monitoring, we also asked whether they would be willing to share this data with doctors, researchers, or fitness apps. We also asked two yes-or-no questions about facial recognition (related to the data type of Face Images), i.e., whether they would use this feature and whether they would allow it to be used on them. The data types we asked about are based on current and anticipated AR and VR glasses data collection features (see support in Sections 2.4.1 and 2.4.2 (b)). While these cover a broad range of possibilities, they were not exhaustive and were kept general enough to be understandable, to avoid distracting or confusing participants with specific details about sensors or cameras, which may also change over time. These questions allowed participants to envision use cases before being prompted with benefits, harms, or questions about data use, which elicited thoughts about possible applications of data types.

If a participant expressed clear comfort or discomfort with the collection or use of a particular type of data, we prompted participants with a pre-written example use case of a benefit or a downside, whichever contrasted with their initial opinion, and we kept track of whether they modified their response to the opposite comfort attitude based on our prompt or maintained their initial stance. After neutral or unfamiliar responses, we provided examples of both benefits and harms. Benefit and harm examples encouraged participants to consider attitudes contrasting with their primary attitude, intentionally probing whether examples could influence their opinion, rather than relying on existing knowledge (or unfamiliarity). In response, participants generally provided conditions for maintaining their original stance or acknowledged the benefit or harm but did not change their overall attitude. For the four instances where participants changed their attitude, we report only the final attitude in our summary of participant attitudes.

We did not prompt participants to discuss privacy concerns and avoided the word "privacy" in our questions, focusing instead on obtaining participants' concerns specific to data collection and use.

We ended the interview with 11 general questions (GQ) about whether certain aspects of AR glasses data collection or data use would make a difference to participants. These aspects included location (GQ1), time of day (GQ2), certain social contexts (GQ3), data subjects (GQ4), data collectors and receivers (GQ5), data storage (GQ6–7), deletion options (GQ8), data transfer options (GQ9), and data collection notifications (GQ10–11). The structured parts of the interview script are included as Appendix A.2.

### 2.5.3 Data Analysis

We used a mix of a priori and emergent coding to code the responses to the 23 questions about data collection and use. We referred to the audio as needed to disambiguate the text and gain insight through prosody, tone, etc. After the interviews, we segmented the interview transcripts into sections, focusing on 23 questions about data collection and use.

**Emergent coding.** Three researchers constructed emergent codes based on their memory of participants' responses, themes in the interviews, as well as codes that would help us label general attitudes: stated positions such as Would Use/Would Not Use and Existence Okay/Existence Not Okay, and Conditional for stated conditions (e.g., "only if I consent to it"). Two researchers then used this initial list to each code two distinct transcripts and refine the codes. These two coders then double-coded all of the interviews using agreed-upon refined emergent codes, splitting the work: each coder was the primary coder for one set of interviews, responsible for coding a set of assigned interview segments, while the other coder was the secondary coder, responsible for reviewing the emergent codes and adding new ones based on their review of the same interview segments. Disagreements, added codes, and code definitions were resolved and clarified through discussion. While we believe that these codes sufficiently capture the major themes we observed in our interviews, we also note that they were the researchers' interpretations of connections between the transcripts and our research questions and that other researchers might generate a different list.

**Coding of attitudes.** To capture participants' attitudes toward AR glasses data collection, we labeled each question's responses with one of three codes: "Comfortable/Support," "Uncomfortable/Oppose," or "Mixed/Conflicted," first assigning the coding of one interview to two researchers. We calculated inter-rater reliability using Cohen's Kappa, which was 0.72, which we considered acceptable to proceed. In addition, we discussed and resolved all disagreements and refined what criteria were used to assign labels for each of the three attitude codes, proceeding without further double coding. We used the refined criteria to separately code the rest of the 20 interviews, with one coder coding 17 interviews and another coding three.

### 2.5.4 Limitations

Our sample of participants is small and not representative of the general U.S. population, Our study purposely selected current users of AR technology, who may have been more comfortable than other people with potential data collection by AR glasses. Thirteen participants used location-based AR games, which may have influenced how comfortable they were with location and image data collection.

We asked interview questions in the same order for every interview, which may have led to an ordering effect. For example, we noticed that some participants' emotional intensity dropped off later in the interview; later replies were more often direct and succinct, such as, "same" or "like I said before."

### 2.5.5 Participant Demographics and Prior AR Experience

Full participant demographics are shown in Appendix A.4. While diverse in terms of gender (10 male, 6 female, two nonbinary, one trans man, one with multiple gender identities, and one who preferred not to respond), most participants identified as white and were under 35 years old. Of the 18 participants who provided income information, eight reported making over $100,000 per year, and four reported making under $30,000. Thirteen of our participants were recruited from Reddit AR mobile gaming communities, seven from other Reddit AR communities, and one from an email listserv (see Section 4.5). Current AR technologies used most often by participants were mobile AR games (n=12), Hololens (n=6), and other AR mobile apps or features (n=3).

## 2.6 Results: Attitudes and Concerns (RQ1)

We first discuss participants' overall attitudes toward current AR technology and future AR glasses data collection and data use (Section 2.6.1). We then discuss participants' concerns about future AR glasses data flows. Some concerns specifically addressed data types from our study questions (Section 2.6.2), while others focused on private spheres, e.g., the home or romantic relationships (Section 2.6.3). Throughout the paper, we use the terminology in Figure 2.1 (from Emami et al.) to refer to number or percent of participants [112].

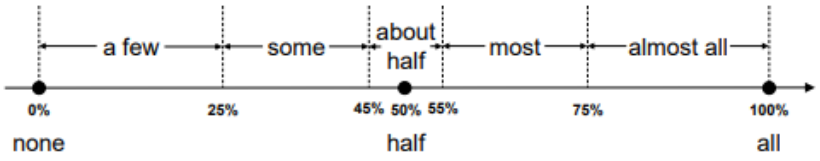


Figure 2.1: The terminology we use to report percentage of participants.

### 2.6.1 Overview

As described in Section 2.5, we asked participants how comfortable they were with current AR technology they used and how comfortable they would be with the collection of 15 data types and five data uses by hypothetical AR glasses. We coded these responses as Comfortable/Support, Mixed/Conflicted, or Uncomfortable/Oppose. Only one participant was uncomfortable with data collection by their current AR device or apps; they said they understood that the data was "very valuable and useful" from a developer and marketing perspective, but that they found it invasive as a user. Some participants were comfortable (n=8) and about half had conflicted feelings (n=11) about data collection by their current AR device or app. One participant's current AR technology was collecting data about others, so we did not ask them this question.

When we asked about future AR glasses, over half of participants, as shown in Figure 2.2, were comfortable with the collection of five data types and with three data uses, and were uncomfortable with or conflicted about ten data types and two data uses. Participants' attitudes ranged

considerably: almost all participants (n=17) were comfortable with location data collection, and fewer than three were comfortable with the collection of face images and bystanders' voiceprints. We present specific concerns about certain data types and uses in Section 2.6.2.

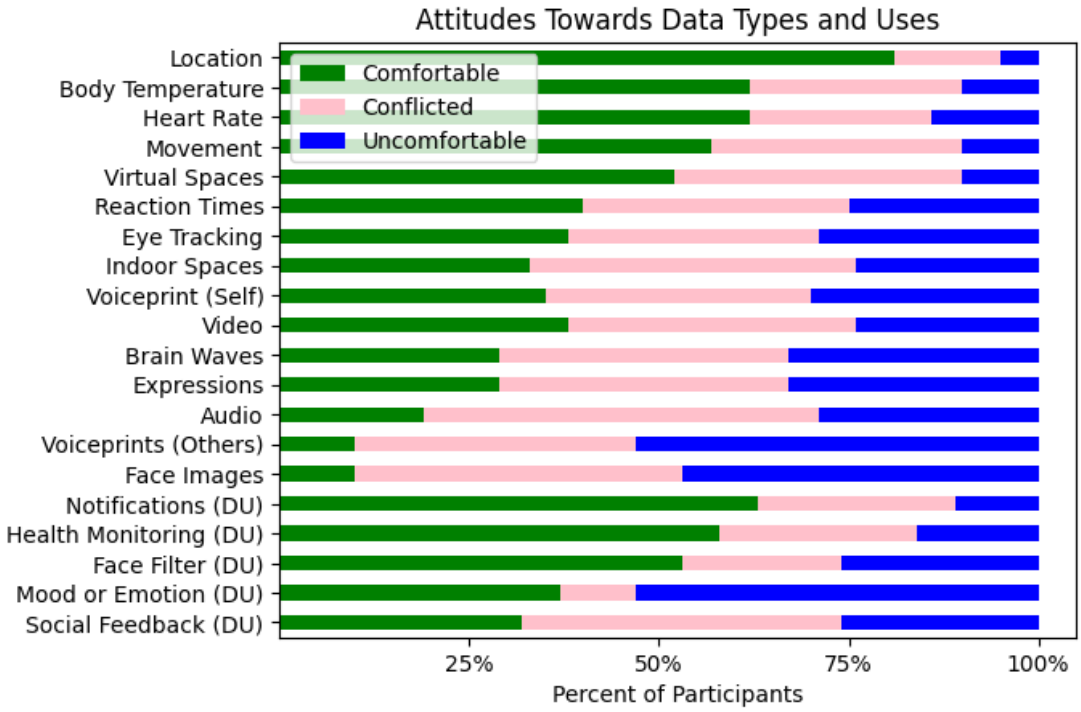


Figure 2.2: Number of participants expressing given attitudes (comfortable, mixed, uncomfortable) regarding the collection of each data type or data use (DU).

Examining these attitudes by participant, we find that 13 participants mostly expressed Comfort/Support, while five showed primarily mixed sentiments (Mixed/Conflicted), and three expressed mostly discomfort (Uncomfortable/Oppose), as shown in Figure 2.3. Almost all participants expressed mixed feelings or discomfort regarding at least nine data types or uses, and all participants expressed either mixed feelings or discomfort in responses for at least three data types or uses (types and uses varied by participant). There were only four instances of participants modifying their comfort-levels, regarding four different data types.

Our coded emergent themes capture recurring concerns and sentiments that shed light on what participants considered to be boundaries, limits, or norms for data collection and use. Fourteen or more participants mentioned the following eight concerns: Recording, Bystanders, Data Use/Purpose, Consent/Opt-in-out, Storage Matters, Context/Situation Matters, Data Protection, and Advertising. Among these, the most frequently mentioned were: Data Use/Purpose, Consent/Opt-in-out, and Storage Matters, being mentioned a total of 71, 63, and 47 times, respectively. All codes and definitions are included in Appendix A.3.

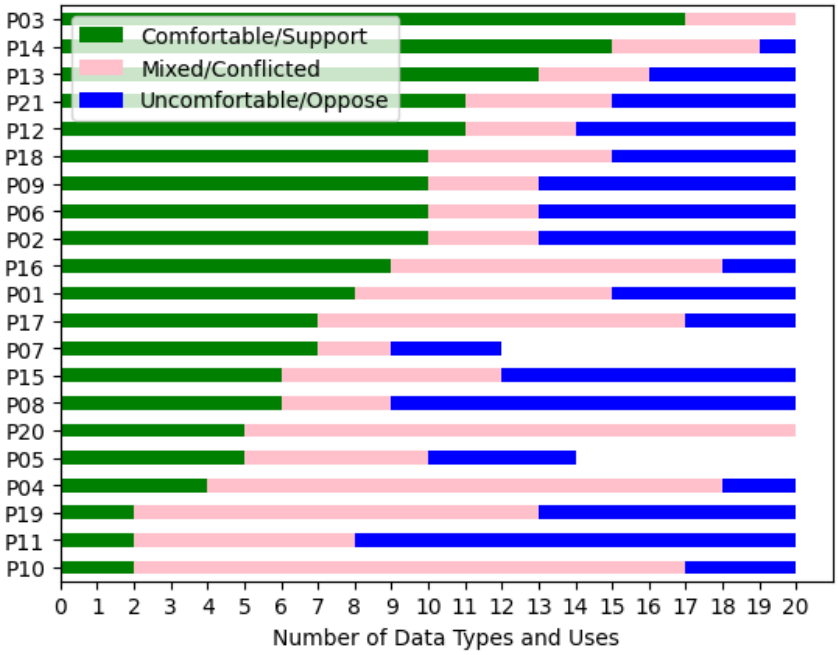


Figure 2.3: Participants' attitudes (comfortable, mixed, uncomfortable) regarding 15 data types and 5 data uses. Note that for some data types, counts do not add up to 20 because we did not ask some questions due to lack of time.

### 2.6.2 Data Types

Participants' attitudes towards data collection varied across data types, with no participant uniformly comfortable or uncomfortable with the collection of all of them. Some participants noted that they were already used to the collection of certain data types, such as location data (n=10) or heart rate data (n=7), for example by mobile phones or fitness trackers. Audio and video data, though used in common technologies like mobile phones and tablets, inspired discomfort based on factors such as recording and storage. We highlight here some data types about which most participants expressed discomfort or conflicted attitudes: Face Images (n=19), Audio (n=17), Video (n=13), Facial Expressions (n=15), Brain Waves (n=15), Mood or Emotions (n=12), Eye Tracking (n=13), Voiceprint (user: n=13; bystander:n=17), and Reaction Times (n=12).

#### 2.6.2 (a) Face Images

Most participants objected to AR glasses using face images for facial recognition on faces detected by the glasses. P8 considered displaying bystanders' personal information (upon recognition) to be "too personal" and unnecessary. P16 said they wouldn't want their face to be "mapped" or photographed and "uploaded to wherever." P5 and P6 suggested that using AR glasses for tasks like processing faces and expressions could cause a decline in natural human abilities:

> *At that point, technology sort of crosses a line where if it's doing a lot of the human interaction for you, I feel like that's something that it's only going to lead to a dark tunnel for us, human nature wise. (P6, Face Images)*

P8 suggested facial recognition would be extraneous, saying, "I don't need the device to tell me if I recognize a person."

**Audio and Video** Participants opposed to audio recording (n=14) or video recording (n=10) were sometimes concerned with AR glasses potentially collecting audio data or recordings of private conversations. P15 showed concern for users whose sensitive conversations could cause them to be "tied back to someone," which could have "legal ramifications for them, social ramifications" (Audio). P4 said they worried about potential harm resulting from someone else possessing audio recordings of them, asking, "What will someone else do about them, to me?" Some participants shared concerns about always-on recording, even if they did not consider their given action or location at the moment of recording to be private per se. P13 gave an example of a prior incident with AR glasses (Snap Spectacles) in which their friends were concerned about being recorded:

> *When I'm trying to wear those glasses and, you know just going to a beach and trying to record something like, my friends will literally get nervous, like, "Hey, are you actually trying to record me?" (P13, Video)*

P7 said they "would feel uncomfortable with any video or photo recording that was not something [they] had specifically pressed a button" to record, conveying a preference to opt-in every time.

### 2.6.2 (b) Facial Expressions and Mood.

Concerns about facial expression data included self-consciousness, involuntary disclosure, and being forced to consider how they present. While some participants suggested that feedback from AR glasses could help them reduce social anxiety, others said it could make them overly self-conscious if their facial expressions "were interpreted as something other than what [they] perceive them to be them about" (P9). P4 said they wanted the "privacy" of "not telling" or disclosing thoughts:

> *Maybe I'm thinking something but don't want to communicate it outwardly, but my face shows it. So when can I have that privacy of not telling someone I'm making a 'That's disgusting, why did you say that?' face [when] I'm saying, 'Oh yeah, that's awesome, thanks for sharing!' out loud. (P4, Facial Expressions)*

While some participants expressed enthusiasm about mood or emotion data helping manage anger or other emotions, about half had concerns. P9 was concerned with becoming "hyper-aware" and "overthinking." A few saw it as unnecessary and overreaching, or said it might exacerbate negative feelings.

> *No one likes to be told, "Stop being angry," and I can see where that might cause escalation. (P4, Mood or Emotions)*

While P6 liked the idea of a mood sensor, they wanted it to be an "indicator" of what they were currently feeling rather than something that "guides you to do something else." Three participants (P14, P10 and P8) considered mood or emotion information to be mental health data, and P14 called for it to be "protected under HIPAA." P8 did not think AR glasses should be tasked with mood evaluation, saying, "It should be best to have a doctor evaluate you for any mood issues."

#### 2.6.2 (c) Brain Waves

While some participants acknowledged potential health benefits of brain wave data collection, most expressed concerns. P17 said it would be inappropriate for consumer devices:

> *There's no reason for that to be in a regular consumer headset. I think that's just outside of the scope of what people buy these things for. (P17, Brain Waves)*

A few suspected harms or invasive applications, with P12 saying that "it feels a little bit like mind reading," and that they "don't want anybody knowing [their] internal state." P7 noted that "brains are incredibly complicated" and suggested that "it's very easy for data to be misused" or misinterpreted, e.g., through "workplace wellness programs." P19 suspected that brain wave data could be used "to track people with some kind of log that can identify your personal brain" based on patterns. P15 suggested it could help create an "advertising monster" to "get [them] to buy more [things]":

> *I wouldn't want, based on what readings are being taken from my brain, [to] then have correlating visual material presented: "Hey we noticed you're depressed all the time, based on your brainwaves here. Try to talk to your doctor about this drug." You know, that's where I get really afraid, is like the advertising monster that AR can enable. (P15, Brain Waves)*

A few participants said they would like to have more information before deciding to use the feature. P18 felt its strangeness or novelty warranted an official company explanation, given the "implicit power in that data and personalized nature of it."

#### 2.6.2 (d) Eye Tracking

While most participants had concerns about eye tracking data collection, a few participants thought it would be necessary for the glasses to function or were resigned to it, noting it was already happening in retail stores (P6). P12 said they would be uncomfortable with eye tracking data being combined with video data in the home. P2 suggested that eye tracking data linked to past experiences could "be negative" because it may evoke "bad experiences" and exacerbate anxiety-related ruminations. P13 preferred that AR glasses not collect recordings for eye tracking data, but rather processed data, e.g., eye gaze vectors.

Some participants expressed concerns about eye tracking data being used for advertising and a few said they did not want ads to fill or obstruct their view. P21 hoped for restrictions or ad blocking features, for example on "when and where" ads could appear, or legal restrictions "on what developers and companies that build applications are allowed to do in terms of advertisements."

#### 2.6.2 (e) Voiceprints (User and Bystanders)

Participants' concerns about voiceprints (distinctive patterns based on a person's voice that can be used to identify, authenticate, or impersonate that person) included storage, bystanders, and impersonation. P20 was concerned about biometric data being leaked if not stored securely and preferred that voiceprints only be used for authentication. P15 suggested having a local voice profile to lock the device and prevent unauthorized users from issuing voice commands:

*I would be okay with them ...collecting enough data to build a local profile on the device where it can decipher my voice from everyone else ...so someone can't do what people did with Google glass—run up and say, Hey Google, search blah blah blah ...and open it in 100 tabs. (P15, User Voiceprint)*

P19 considered voiceprint data collection to be "more about access to your microphone," suggesting that any audio data collection could potentially capture voiceprints. P10 suggested that recording laws should apply to voiceprint data. Some (n=8) participants expressed concerns about bystanders (see also Section 2.7.2). P16 suggested having a "guest mode" or "visitor mode" to remove personalized analysis of bystanders' voice patterns. A few participants were also concerned about impersonations or "deep fakes," and suggested imposters could use voiceprints to demand money from family members, falsely authenticate, produce a fake conversation they never had, and use voices for advertising.

#### 2.6.2 (f) Reaction Times

Three participants suggested that reaction times could reveal health changes, such as worsening conditions, but they felt differently about this possibility: while P18 found it helpful, P20 said they weren't sure, and P2 said they would not want this data to be revealed or shared by the AR glasses. Most participants were especially uncomfortable with reaction time data being collected by insurance companies after we gave an example of them increasing premiums after observing slow reaction times on breaking, with P3 saying, "Even the best of drivers would never agree to something like that." P16 noted that reaction time data could already be collected by car insurance telematics systems.

### 2.6.3 Private Spheres

Participants expressed concerns about data being collected in certain locations or social situations, like activities and relationships.

#### 2.6.3 (a) Private Places or Physical Locations

Specifying spaces where they would want to limit or disable data collection, about half of participants (n=12) mentioned the home, five mentioned the workplace or office (see Section 2.8.4), and a few mentioned the following: doctor's office, car, government facilities, funeral places, LGBTQ+ meetings, political meetings, restaurants, church, and parties.

**Home.** A few participants discussed their privacy concerns about the home. P12 consistently had mixed feelings about sharing data, depending on whether they were in public or at home:

*Within the home, it's kind of like a sacred place where you can be weird and goofy, ...and nobody should be able to see that but you. (P12, Video)*

A few participants focused on specific rooms or activities within the home. P5 did not want AR glasses collecting video data when in the shower or on the toilet, or when "engaged in intimate actions with [their] partner." P18 felt uncomfortable with the idea of AR Glasses sending reminder notifications to clean the toilet. P13 said they would rather take their glasses off while in

the home, given all the information that could be collected about them and their child. P4, P5, and P13 expressed concerns about security risks of AR glasses collecting maps of indoor spaces, e.g., burglars, with P4 suggesting the data could disclose where users keep valuables or where their children sleep.

**Privacy in Virtual Spaces** Most participants (n=13) expressed discomfort or mixed feelings about data collection in virtual spaces, and three (P7, P13, P21) had concerns about potential identity disclosure, suggesting that virtual spaces allow a freedom of behavior or interaction that might be infringed upon if identities were revealed.

> *Some people, they ... completely become like another different person when entering a game world, so they can be free for themselves, they may go out of boundary a little bit. (P13, Virtual Spaces)*

If such people were being "tracked," P13 suggested, they might not feel as free. P7 mentioned the risk of malicious actors identifying and targeting members of identity-based virtual spaces, for example, in a social VR space that "caters to trans people."

### 2.6.3 (b) Private Activities and Relationships

Some concerns were not about spaces but rather activities or relationships, such as recording, family relationships, and romantic relationships.

**Family Relationships** Some participants placed limits or expressed discomfort regarding data collected about their family, such as conversations within the home with family, or where children sleep, as noted above. P14 considered how lie detection using facial expression data might influence parent-child power dynamics:

> *Could you imagine how powerful parents could be if they could facially detect their children when they're lying to them? (P14, Facial Expressions)*

**Romantic Relationships.** Some participants noted concerns relating to romantic situations, such as dating or intimacy. P16 gave an example regarding disclosure of personal information on a date:

> *I don't want the Black Mirror episode where a couple is going on a date and the dude is getting all her information just by looking at her. That's creepy land. (P16, Face Images)*

Two participants (P4, P5) expressed not wanting the glasses to collect data while they were engaged in physical intimacy. P4 specified a time span that should be off-limits:

> *Not to be crass, but I don't want it to have it from these hours, where I know I was having sex or something, ...unless it has some benefit, because I [would] know I'm having like a heart issue, and they need to see it during high activity. (P4 Health Monitoring)*

## 2.7 Results: Data Actors and Subjects (RQ2)

Participants expressed concerns about data collectors, receivers, and subjects (Sections 2.7.1–2.7.2), as well as concerns about potential harms to people in vulnerable situations or populations (Section 2.7.3).

### 2.7.1 Data Collectors and Receivers

Most participants (n=16) expressed reservations about data collectors and receivers, and all 13 participants who were asked whether they would like to know who or what companies would have access to AR glasses data about them (Q5) said yes. Potential collectors and receivers about whom we prompted participants included the AR glasses company, employers, insurance companies, advertisers, doctors, researchers, and fitness apps. Participants raised concerns about unknown receivers, law enforcement (see details in Section 2.7.3), burglars, home attackers, kidnappers, hackers, someone you're trying to avoid, domestic violence abusers, and stalkers.

#### 2.7.1 (a) AR Glasses Company

A few participants expressed opinions regarding the AR glasses company. For example, P8 considered Apple to be "secure" and Amazon to be "a trusted brand," adding, "If it's a brand I'm unfamiliar with I wouldn't want them listening to my conversation" (Audio). P14 said they "wouldn't trust Facebook at all" because of their "spotty history" with privacy, but said they "might trust Google a little more" (Indoor Spaces). P19 expressed skepticism about AR companies deleting data:

> *Even if you delete your Facebook data, they'll just start a new tracker profile for you. ...We don't have any real regulations, at least in the US, about whether that data is truly deleted. I would want to see some kind of regulation about that first, before I would even trust the feature. (P19, Q8: Deletion)*

P15 said their comfort transferring data across device brands would depend on if they agreed with each brand's storage policies (Q9).

#### 2.7.1 (b) Employers

Some participants were concerned about data being sent to employers. P7 was concerned about gestural data potentially being "mis-use[d]" by employers to flag people as suspicious (Movement). Similarly, P4 suggested consequences for perceived "aggressive" movements:

> *If I'm behaving a certain way at work, and this starts triggering something, ...like, oh, she must be aggressive, the way she's waving her hands. Now we're flagging your HR file. (P4, Movement)*

P12 appeared to note uneven benefits:

> *An employer, looking at the reaction time of all their employees and saying this person's really slow, they're getting fired, right, it seems to benefit external third parties more than the [data subject] (P12, Reaction Times)*

P11 suggested that employers' access to social feedback and mood or emotion data could be "bad for me, employee, but good for that business, because you know, I'm not improving their business." P14 was concerned about "employers reaching too far into the personal lives of employees" and suggested that employees might be pressured into sharing information (Social Feedback).

#### 2.7.1 (c) Insurance Companies

Even without being prompted by a negative example (Reaction Times), some participants expressed concerns about medical or life insurance companies increasing rates or denying coverage based on AR glasses data.

> *If the device decides that I am having seizures ...does that mean like okay now I have pre-existing conditions and I can't get insurance? (P4, Health Monitoring)*

P4 said they would not want AR glasses to be able to automatically "report" health conditions to doctors and insurance companies.

#### 2.7.1 (d) Health Monitoring Data Receivers

About half of participants said they would share health monitoring data with doctors or other healthcare providers, researchers, and fitness apps. Some participants said they would share information with doctors if they were experiencing a serious health issue. A few expected this data to be protected by U.S. HIPAA law, and others noted they would not want it to be shared with insurance companies or a nationwide hospital system. When asked about sharing with researchers, participants were concerned about consent, the purpose of the research, how the data would be used, funding resources, who the researchers were, and secure storage of the data. Participants' concerns about fitness apps included whether they trusted the app or platform and if they could control who uses the data.

#### 2.7.1 (e) Advertisers

Most participants (n=14) anticipated advertisers becoming involved in their data flows, and a few were concerned that advertising could be excessive or annoying.

> *I wouldn't want targeted advertising in a virtual space, because I don't see the point. Why would you interrupt my virtual experience with something from the real world? (P9, Virtual Spaces)*

P5 said they did not want advertisers "to know what excites [them]" and "get more effective," or gain a deeper understanding about them, because they already found ads to be "intrusive" (Brain Waves). P10 expressed cynicism about advertisers' motivations to learn what would make users "respond with a knee jerk reaction" and exploit it to "get people to impulse buy something or get them irrationally angry about issues" (Reaction Times). P2 recalled feeling conflicted about making purchases in their mobile AR game:

> *Lately ...should I be buying random cosmetics in Ingress? Probably not, but I do it, and if they're offering more, then I might consider buying more stuff. (P2, Virtual Spaces)*

P2 suggested such pressure could lead to "addiction," comparing it to Amazon's purchase suggestions.

### 2.7.2 Data Subjects

When perceiving themselves as the data subject, some participants expressed concerns about disclosing personal information and hoped that AR glasses data would be anonymized or used in aggregate across groups of users, rather than associated with a personalized data profile based on inputs such as "eating habits, shopping habits, [or] health activity" (P4, Brain Waves). A few specified desiring or expecting data about themselves to be encrypted.

Almost all participants (n=18) expressed concern about bystanders' data being collected or used, mostly for the data types of face images, facial expressions, and bystanders' voiceprints, with eight participants suggesting that bystanders should be able to consent to data collection and a few suggesting that faces should be blurred, features should not work on other people, and facial recognition should be illegal. A few were concerned about the possibility of bystanders' data being collected by law enforcement, government, or security authorities (discussed in Section 2.7.3). P4 said they would be "worried" about "profiles that are being created about who I see and I interact with" (Face Images). Nevertheless, eight of 17 participants said they would use facial recognition on others, while only four of 16 said they would allow it to be used on them. Thus, more participants objected to being a subject of facial recognition than to being a collector. P14 thought data about other peoples' facial expressions could help them make a sale:

> *It might be useful for me to go back and see, if I was trying to make a sale ... what people's reactions were to how I phrased various specific things. (P12, Facial Expressions)*

P14 and P16 suggested facial expression analysis could be used to detect lying in negotiations, depositions, and parenting.

### 2.7.3 Vulnerable Situations or Groups

We highlight below concerns about potential harms for people in vulnerable situations or groups. Some potential traits or situations mentioned by participants include: disability, non-standard accent or self-presentation, associating with or being near someone sought by law enforcement, dissenting or protesting, marginalized gender or sexual identity, and being a victim of stalking, abuse or theft.

**Normative Biases.** Some participants (n=12) suggested that discrimination could result from normative biases built into AR glasses features. For example, P7 was concerned about facial expression analysis on people who might not express themselves in normative ways, including autistic and other neurodiverse people:

> *Video analysis programs for use in hiring and video interviews, that is a real problem, that people who do not express their emotions typically can be flagged as untrustworthy or suspicious. I'm autistic. I don't necessarily express myself the same way*

*that neurotypical people do. If my phone was paying attention to my facial expressions to try to judge my mood and respond in different ways, I feel like it just wouldn't work very well. (P7, Facial Expressions)*

P20 imagined a social feedback feature negatively assessing their non-standard American English accent from "rural Appalachia":

*If the glasses are telling me that my accent is poor and that I need to retrain myself how to speak, that, I would have a bit of a bigger issue with. Having grown up in rural Appalachia, I can attest to the fact that an accent does not determine how intelligent or capable another person is. (P20, Social Feedback)*

Such concerns capture some potential AR glasses use cases that exclude minority or non-standard populations.

**Criminal or Political Punishment.** A few participants were concerned about authorities receiving AR glasses data and using it to target or harm people wanted for criminal or political activity. P4 was worried that face image data could cause them to set off a crime alert simply by observing a bystander who was sought by police, potentially "creating some sort of risk level" and profile for them (P4), as well as potentially involving them in the capture of an innocent person. P5 expressed concern that law enforcement's use of faulty facial recognition would result in "unfairly targeting the people that it's least able to recognize."

A few participants were concerned about AR glasses exposing users' or bystanders' political views, gender or sexual identity. P14 was concerned that attendees of meetings based on LGBT+ identity or political views could be identified via voice recordings:

*I'm most nervous about, say, people who are LGBT plus and not out of the closet getting recorded at meetings. ...Or political meetings, like everyone knows what happens in certain countries. (P14, Bystander Voiceprint)*

P20 suggested face image data of "political activists" could be collected at protests by "public security individuals" (P20, Face Images).

**Personal Threat** Some participants (n=11) also expressed concerns about safety (their own or others'). Five participants mentioned the threat of stalking, and three were concerned about burglars. P7 expressed concern about inadvertently distributing information about a friend who had been stalked before:

*I have friends who don't like having their pictures taken at all and don't want their pictures going on social media ...İf I had glasses that did facial recognition, it would be kind of a betrayal of trust of those people. (P7, Face images)*

P14 evoked data risks faced by abuse victims in shelters:

*[If someone] has one of these inside of a domestic violence shelter and isn't turning it off, that kind of thing, ...the custodian of that data ends up being very important in that situation. (P14 Indoor Spaces)*

This suggests that AR glasses disclosing the location and identities of people in the shelter could endanger the user and others.

## 2.8 Results: Desired Privacy Principles (RQ3)

We present some privacy principles that participants expected or wished to apply to AR glasses data collection: user control, notice and consent, need-only use, and data protection. All 21 participants placed conditions on data collection or use, with conditions being noted across all categories a total of 125 times. The most common ones were notice (n=13), consent (n=19), and storage (n=16).

### 2.8.1 Individual User Control Over Data Flows

All participants expressed a wish to be able to customize or limit data collection, i.e., to enable or disable it, delete data, and control where and how data is stored. Most participants (n=16) wanted the ability to turn off certain features, and about half (n=10) wanted certain features to work only if initiated by the user.

> *If I couldn't stop it from recording everything, yeah, that's a hard no. (P10, Audio)*

P3 wished to control timing of data flows, to know "when" and "what data" was being sent to doctors and to not send "a constant stream of information to them" (Health Monitoring). P12 wanted granular control over "which data streams were being used" and was against brain wave data collection but would accept body temperature and heart rate data collection (Health Monitoring).

**Ability to Delete** We asked nine participants whether the ability to delete their data made a difference (Q9), and eight said yes.

> *As the creator of that data, you should also have custody over that data and your right to have custody over that is also your right to delete it. (P20, Q9)*

Some participants were skeptical that they would even have this option, suspecting that their data would exist somewhere else anyway. Participants considered deletion useful for when they stopped using the device, wished to be forgotten (e.g., delete their social media account), and to delete data they did not find useful.

**Storage.** Most participants (n=16) also expressed concerns or reservations about storage and retention of the data. Six participants noted that they would rather have data be used temporarily and not saved or stored, with a few mentioning their home IoT devices having such options for ephemeral or temporary storage. When asked specifically about storage (Q6), 11 of 12 participants expressed a preference for data to be only stored locally on the device, not on a remote server ("in the cloud"), with a few participants being concerned about data breaches. P12 said they would have "low confidence that anything that happens on the cloud will stay truly private in perpetuity. P6 and P15 conditioned their comfort with cloud storage on encryption and hashing.

### 2.8.2 Notice and Consent

About half of participants (n=13) expressed a wish for notices. When asked about how they would want to be informed about who has access to their data (Q5), participants suggested being

informed through an initial walk-through or notification, a settings menu, email, privacy documents, semi-regular reports, a website, or mobile and desktop apps. P15 wished for granular explanations about how health monitoring data would be used. Almost all participants (n=19) said they expected or desired a choice, i.e., to provide consent, opt in or out of data collection or certain features, or rescind or modify their consent, mentioning consent a total of 77 times for all data types and four data uses (all except Face Filter).

### 2.8.3 Need-only Use

A few participants acknowledged that some AR glasses data collection might be "necessary," and thus that users could sometimes lack meaningful choices, given no alternative except to not use the device. P19 suggested that users would not be able to opt out of eye tracking features, since disabling them might "make the device basically inoperable" (Eye Tracking). P4 objected to potential uses beyond the improving an application:

> *I would be okay with that, unless it's trying to create a behavior profile of me. ...If it's a more generic sense, to understand how players or users of the system interact, to increase usability or productivity of the feature or device, that's fine. (P4, Virtual Spaces)*

Other participants also conditioned their comfort on whether data collection was necessary for the functioning of the feature or device.

### 2.8.4 Data Protection

Ten participants raised concerns about sensitive data content, across nine data types and two data uses, such as inferred income level, personal identity, sensitive conversations (with friends, family, romantic partners, coworkers, or in multi-player games), data profiles, certain images, and information collected while sleeping. Some participants suggested that legal protections should apply to certain data. Four participants felt biometric or health data should be protected by health privacy laws, e.g., HIPAA (U.S.):

> *There's already so many protections around medical data, I'm not too concerned about that. (P10, Health Monitoring)*

All 21 participants had concerns about recordings, for 12 data types and one data use. Three participants (P5, P7, P10) discussed potential unacceptability of recording, evoking U.S. state laws against recording without permission, which apply in what are known as All-Party or Two-Party consent states. A few participants expressed concerns about AR glasses violating confidentiality norms, such as intellectual property (IP) rules. P13 suggested a company's strategies or product information could be leaked.

> *Imagine another company trying to get ...information by monitoring what employees are talking about. ... That could be a disaster. It's literally a breach. (P13, Social Feedback)*

For this reason, P13 suggested that information collected in office settings be stored locally on the device "without being analyzed by a third party." P12 suggested AR glasses collecting indoor

spaces data might violate a workplace policy by "giving away trade secrets." P4 raised concerns about ownership of voiceprints:

> *As long as I still own the rights. Like I don't want to go into a store and start hearing voiceover things in my voice. (P4, User Voiceprint)*

They opined that "a recording of Abraham Lincoln's voice …could be really cool" but wondered about the "rights" to Lincoln's voiceprint.

## 2.9 Discussion

Our study reveals many diverse concerns and desired privacy standards for future AR glasses. Some of our findings are specific to AR glasses, due to the mobile form factor of a head-mounted display, the integration and combination of sensors (corresponding to our data types), and the potential for instantaneous analysis of these sensor inputs while moving in and engaging with the external environment. Here we bring together privacy concerns spanning different aspects of AR glasses data collection using a subset of Solove's taxonomy of privacy harms. We also discuss the context-relevant nature of participants' concerns, consider the privacy implications of potential mitigations for concerns about vulnerable situations or groups, and compare our findings to prior work. Finally, we make recommendations for privacy legislation and AR glasses design.

### 2.9.1 Privacy Harms

We apply Solove's privacy harms taxonomy [383] to the concerns raised by participants, noting which problems from the taxonomy's four groups the participants discussed. In comparison to features of widely available devices like mobile phones, several of these could be relatively unique to AR glasses, such as combined video feed, input from multiple sensors, and network connectivity, as well as features like reaction times, brain wave data, facial expression analysis, and overlain advertisements.

**Information collection.** First, participants raised concerns about potentially harmful information collection, specifically the problem of surveillance, such as recording private conversations or activities, tracking brain wave data and correlated visual material, and employers monitoring reaction times of their employees.

**Information processing.** Participants also expressed concerns about all of Solove's information processing problems: aggregation, identification, insecurity, secondary use, and exclusion. Given the multiple sensors of AR glasses and its potential network connectivity, participants suggested various pieces of data about them might be aggregated, such as external sensor data (e.g., video or audio) and biometric data, which could be used to make inferences about users. Identification, or "linking information to particular individuals," was also a concern, especially in virtual spaces or via facial recognition. Regarding storage, some participants suspected problems of insecurity, which Solove defines as "carelessness in protecting stored information from leaks and improper access," expressing a wish for data to be stored temporarily, be deletable, and also

to be stored on-device only. Secondary use ("the use of information collected for one purpose for a different purpose without the data subject's consent") came up several times, as participants expressed concerns about health monitoring data being sent to insurance companies who might deny them insurance based on pre-existing conditions, brain wave data being used to recommend medical treatments or products, and recordings being used by third parties. Exclusion, or failure to inform the data subject about data collected or used about them and to involve them in its handling and use, was a problem raised about bystanders as well as about users whose data might be collected by unknown receivers and used for unknown purposes. Many participants wished to know who would have access to their data and how it would be used.

**Information dissemination.** Participants were concerned about problems related to information dissemination, including the problems of breach of confidentiality, disclosure, exposure, appropriation, and distortion. Breach of confidentiality ("breaking a promise to keep a person's information confidential") was raised as a potential problem in work settings, where AR glasses might cause proprietary information to be leaked. Disclosure ("revelation of truthful information about a person that impacts the way others judge her character") was a concern regarding marginalized identities, such as LGBTQ status, disclosure of negative feelings using facial expression analysis, as well as health conditions, such as worsening health. Some worried about potentially pervasive AR glasses data collection resulting in exposure ("revealing another's nudity, grief, or bodily functions") of activities they considered private, such as bathroom use or physical intimacy. Participants concerned about appropriation ("the use of the data subject's identity to serve the aims and interests of another") suggested that bystander data and voiceprints could be appropriated to serve the glasses user or corporate interests. The problem of distortion ("dissemination of false or misleading information about individuals") arose when discussing data that could be misinterpreted and used by employers, such as brain wave, movement, or facial expression data.

**Invasion.** Some participants expressed concern about privacy invasions such as intrusion (disturbance of tranquility), fearing that AR glasses would pester them with advertisements or personalized recommendations, including behavioral modification suggestions, or wishing to limit data collection in private spaces or contexts.

### 2.9.2 Context-Dependent Privacy Considerations

The range of privacy contexts and concerns provided by participants, even with a small group of people, suggests that it would be difficult to design satisfactory static and predetermined privacy options. While some participants invoked the private/public dichotomy, dividing privacy spheres into two contexts [289, 427], as a basis for constraining data flows, others provided contexts that do not fit into a binary private/public separation of places or activities, such as virtual spaces, the act of recording, or sensitive conversations. Given such contextual dependencies, we encourage researchers and AR professionals to apply the privacy framework of contextual integrity [289, 290] to study AR glasses privacy concerns in specific contexts, to better explore factors such as stakeholders and sociocultural norms, and to discover and articulate information

norms, which Nissenbaum defines using the elements of data types, senders, receivers, subjects, and transmission principles—some of which coincide with topics of concern in our study.

While a norms-based approach to privacy can be useful for establishing baseline online privacy norms rooted in physical social life, Nissenbaum acknowledges that such a framework "appears to provide no buffer against insidious shifts in practice that ultimately gain acceptance as 'normal.'" In our study, participants sometimes expressed discomfort with entrenched data practices, such as advertisers receiving their data, employee monitoring, always-on functionality, and recording of bystanders. Yet, they varied in expressing resignation or a desire for alternatives. A way to avoid further entrenching unwarranted or "tyrannical" normative practices is by comparing these entrenched practices "against novel alternatives or competing practices on the basis of how effective each is in supporting, achieving, or promoting relevant contextual values" [289]. Thus, future work could explore establishing novel alternatives to entrenched norms that better embody relevant contextual values. Articulating what these contextual values are is also a space for future work, since values can fluctuate in online settings based on factors such as who has access to information, who the information is intended for, and technological privacy affordances [258].

### 2.9.3 Potential Exclusion or Discrimination

Some participants raised concerns about AR glasses data collection potentially resulting in exclusion, marginalization, or discrimination, e.g., features that apply normative biases to users with disabilities or nonstandard accents could alienate or marginalize them, especially in the contexts of hiring interviews or social feedback. Future work should explore potential mitigations for such discrimination and their privacy implications. For example, if mitigations are developed to detect or to take as input factors such as disability or dialect, such that the product could adapt to these factors, what are the potential privacy protections that could be placed to prevent this data from being further distributed or used? Also, if AR glasses are able to detect medical conditions and social, linguistic, or demographic information about users, designers and researchers should consider potential harms of unintentional disclosure or false positives of detected features.

Our findings also suggest political and humanitarian implications for AR glasses data policies: dissenters, protesters, LGBT+ people, and innocent suspects were given as examples of people who might be at risk of persecution or punishment if AR glasses data were to be used against them in certain contexts. Additionally, a few participants evoked the norm of anonymity in virtual contexts that permits a certain freedom of behavior or interaction and suggested that mandating disclosure of or exposing identifying information could pose privacy and security risks to people in vulnerable situations or groups. Yet, designers must also contend with the potential for exclusionary norms or harassment in such spaces (see Section 2.4).

### 2.9.4 Comparing Findings to Prior Work

As in prior work on AR glasses privacy concerns, the purpose of data collection or use was a major concern for participants [206], as was recording, with some participants evoking laws against recordings [93, 206, 224]. Similar to findings in Koelle et al.'s work on social acceptability of data glasses, participants responded differently based on whether they considered themselves the user or the subject (in our case, of facial recognition) [206], as well as whether they made the choice

to provide information or not [351]. Unlike most prior work, our study provides an analysis of the privacy concepts or problems evoked by participants. Our focus on data collection allowed participants to provide concerns about specific data types and articulate objections to particular data flows that were not specified in prior work that specified privacy concepts [306].

### 2.9.5 Recommendations

**Design** Participants' privacy wishes and expectations vary significantly, even among 21 participants. Designing for millions of people will surely invite even more complex considerations. Additionally, some challenges (e.g., overlaid advertisements, expression analysis) are new enough that we believe substantial additional research is needed before we can develop designs to address them. Participants' varying levels of comfort based on diverse factors suggest that AR professionals should create flexible privacy choices to meet complex privacy considerations, such as customizable user controls over data flows and the ability to opt out of data collection. There is ample opportunity to provide AR glasses users with choices about data collection, use, and storage, for example, as proposed in prior work [104, 181]. We also recommend designing for adaptive privacy considerations, based on dynamic factors such as the varied privacy needs of people present in the same physical location or virtual context.

Most participants wished to receive detailed information about how their information is used, who receives it, and where it is stored. Many AR/VR glasses and applications already require personally identifiable data for functionality [201]. In such situations, users may not be granted means to control or regulate certain data flows, but notice and consent mechanisms should provide transparency regarding data collection, use, sharing, and storage. Harborth et al. found that contextualized justifications for mobile AR app permissions reduce privacy concerns and increase the willingness to grant permissions [144]. This suggests that contextualizing data collection and use by providing details about intended functionalities and data access could assuage user concerns about opaque or nefarious data flows. However, being transparent about excessive data collection or violations of privacy norms may not necessarily reduce privacy concerns, and managing the intensity and frequency of notifications is an ongoing issue in privacy notice and consent interfaces.

**Legal Protections** We anticipate invasive products that test users' privacy boundaries. Most participants expected or imagined scenarios in which their data was collected or accessed by private actors such as the AR company, advertisers, and their employers. Some expressed discomfort regarding potentially opaque data use and data storage practices. A few participants expected health-related data and recordings to be protected under HIPAA (which only covers data used in a healthcare context) or All-Party recording laws (which vary by state and may not protect biometric information). Recent litigation under laws such as Illinois's Biometric Information Privacy Act has suggested that regulation can be effective in moving companies to develop options for transparency and consent [46, 329]. Without robust privacy laws, production of knowledge and recommendations about potential users' concerns and limits is not likely to result in better protection for consumers [109, 226, 318]. Our study confirms the multi-faceted nature of privacy, with participants expressing concerns that cannot be addressed by one-dimensional definitions of

privacy. We therefore recommend that policy makers enact regulations for data collection, data protection, user control and disclosure, as well as laws that, like HIPAA, enforce privacy norms for data flows in settings where privacy is socially mandated.

## 2.10 Conclusion

Our exploratory study analyzed data from 21 interviews regarding data collection and use by hypothetical future AR glasses. We consider participants' privacy concerns and desired privacy principles, which are varied and context-dependent. We connect these results to privacy concepts and call for multifaceted solutions.

# Chapter 3

# AI-enabled Analysis of Voice Data for Decision Making

This chapter was adapted from my paper currently in submission:

> Andrea Gallardo, Lily Klucinec, Lujo Bauer and Lorrie Cranor. 2025. "Attitudes Towards the Use of AI-enabled Voice Technologies for Decision Making Among Southern U.S. English Speakers."

## 3.1 Overview

In this study, we consider the emerging voice-specific applications of AI in four high-stakes decision-making use cases: hiring interviews, job performance evaluations based on presentations, college admissions interviews, and grading of pronunciation on English final exams. Through a qualitative survey, we documented the attitudes of 111 American English speakers from the southern United States towards AI-enabled analysis of the voice data of data subjects in these four use cases. We recruited American southerners to invite potential insights and opinions on how speakers of non-standard dialects might fare in algorithmic evaluations based on speech, as the rise of algorithmic decision-making in workplaces and schools has drawn attention to potential benefits and harms for job applicants, employees, and students.

Our purposive sampling yielded directly relevant responses that shed light on the experiences of speakers of non-standard dialects and some potential social and policy consequences of AI voice analysis in high-stakes contexts. Especially relevant were insights from participants who identified as speakers of the Appalachian dialect, who described difficulties being understood by technology, social bias and stigma based on their way of speaking, as well as the joy and charm inspired by their dialect. Such reflections offer nuance and culturally relevant insights into such harms as AI-enabled discrimination, the potential cultural loss of regional dialects, and unfair outcomes resulting from technical errors.

This study relates to the thesis through its consideration of a particular population when evaluating the potential harms of emerging technical applications. In considering AI-enabled analysis of voice data in high-stakes contexts, we elicited potential benefits and harms from individuals

who speak or are likely to be familiar with non-standard dialects, i.e., input that may result in skewed results from algorithmic evaluation. In this work and our most recent study on translation technologies, we focus on technologies that process natural language and consider the context-sensitive implications of using these technologies in high-stakes settings.

## 3.2 Abstract

The rise of algorithmic decision-making in workplaces and schools has drawn attention to potential benefits and harms for stakeholders. Additionally, uneven performance of language technologies for non-standard or underrepresented dialects has resulted in calls to engage speakers of those dialects. Our survey study explores the attitudes of 111 American English speakers from the southern U.S. towards voice-specific applications of AI in four high-stakes decision-making use cases: evaluation of candidates in hiring and college admissions interviews, employee performance evaluations based on presentations, and grading of students' pronunciation on English final exams. While participants acknowledged the relevance of metrics such as pronunciation and communication skills to employers and schools, they conveyed concerns about biased evaluations that negatively or inaccurately evaluate certain dialects or speech types, often expressing a preference for human involvement during decision-making. We recommend involving non-standard dialect speakers in language technology design to help promote and preserve language diversity, conduct user-centered AI auditing, and help align AI systems' evaluations of data with the values and rights of stakeholders and with the functional social norms of interviews and performance evaluations.

## 3.3 Introduction

The use of AI-enabled tools to evaluate people in hiring [43, 55, 74, 88, 277, 311], job performance evaluations [189, 331], exam grading [270, 430] and college admissions [70, 183, 205] processes is an increasingly popular technological affordance for decision makers. Products marketed to managers, teachers, and administrators claim to use statistical modeling to evaluate how well data subjects meet performance metrics or eligibility standards and algorithmically filter, rank, or rate them. Some products also make overreaching or pseudo-scientific claims of assessing complex human traits of job applicants based on job application videos [330, 379]. While using AI-enabled analysis of voice data to assess human standards of pronunciation [32, 87, 111] and communication skills [132, 396] may be relevant to education and employment contexts, and while such tools could help mitigate human biases affecting a company or school's decisions [59, 272], prior work has shown that algorithmic evaluation of human traits such as personality or hireability can be unstable or biased [147, 346, 347], pseudo-scientific [129, 143, 271, 330], and lead to unfair outcomes [117].

Our work focuses on potential societal harms resulting from the use of AI analysis of voice data for high-stakes evaluations, which may produce erroneous or biased outputs based on factors such as non-standard speech. Indeed, recent work demonstrates the uneven performance of automatic speech recognition (ASR) technologies, with some performing worse for non-standard

English dialects or accents [251, 366].

In the U.S., sociolinguistic differences have a history of resulting in discrimination in education and employment contexts. Thus, erroneous or biased algorithmic evaluation of voice data could potentially worsen social discriminatory patterns facing speakers of non-standard dialects. We therefore intentionally engaged with speakers of non-standard and low-resource dialects who may be particularly affected by AI-enabled analysis of voice data: Southern U.S. English speakers. We focus on this population because the Southern U.S. is a region where many people use Southern U.S. English dialects and linguistic features that are considered non-standard or low-resource in the context of common ASR technologies such as voice assistants (VAs) [211, 217].

Through our survey with 111 people from the Southern U.S. who self-identified as speakers of Standard American English (SAE), Southern U.S. English (Southern), African American English (AAE), or Appalachian English (Appalachian), we investigate participants' attitudes towards voice-based AI analysis, using vignettes depicting four high-stakes decision-making use cases: hiring interviews, job performance evaluations based on presentations, college admissions interviews, and pronunciation grading on English final exams. Specifically, our research questions are:

- RQ1: How acceptable do participants find:
  - (a) the general use of AI to assist with decision-making in four high-stakes use cases?
  - (b) the use of a voice-specific application of AI in these use cases?
  - (c) voice data collection by data collectors potentially involved in the four use cases?
  - (d) inferences made by voice assistants about speaker attributes?
- RQ2: What benefits and harms do participants anticipate regarding the use of voice-specific applications of AI in high-stakes decision-making contexts?
- RQ3: How do participants relate their attitudes towards speech-based evaluations to their experiences with non-standard dialects, accents and speech?

We asked participants to rate the acceptability of the following: general and voice-specific applications of AI in the four aforementioned use cases, data collectors potentially involved in the four use cases, and speaker inferences based on voice data, such as speaker gender or dialect. We also solicited written responses addressing each use case to capture overall attitudes and perceived benefits and harms associated with the voice-specific applications of AI.

We provide a detailed qualitative exploration of participants' stated attitudes, concerns, and experiences. While some participants highlighted positive perceptions of AI use cases, such as increased objectivity or efficiency in decision-making processes, a majority expressed concerns about potential bias or discrimination resulting from systems that cannot understand or negatively evaluate certain dialects or accents. Participants also conveyed their preference for humans to review algorithmic output or to be solely responsible for evaluations and decision-making. Finally, some participants volunteered positive and negative experiences related to their non-standard dialect, accent, or speech, providing insights that both support and oppose the reinforcement of standard language use by AI technologies.

Our findings shed light on the risks of systematizing linguistic bias and perpetuating discrimination based on speech. We recommend future work that explores how to better align AI systems with societal standards for equal opportunity in employment and education contexts, as well as

participatory research engaging speakers of non-standard dialects, user-centered AI auditing, and privacy regulations.

## 3.4 Background and Related Work

We provide some background information and prior work on AI-enabled and algorithmic evaluations in high-risk contexts, sociotechnical harms of algorithmic systems, disparate performance of speech recognition technologies across dialects, and marginalized American English dialects.

### 3.4.1 High-Risk AI in Education and Employment Contexts

Previous work in a variety of fields has demonstrated the employment [65, 142] and educational [31, 168, 393] challenges faced by those in diverse racial, gender, and socioeconomic groups. However, the growing use of AI in professional contexts [117, 231, 331], with its potentially pseudo-scientific claims to assess complex human traits [379] and its unstable or biased evaluations [25, 346, 347], has resulted in an increase in calls for enforcement of anti-discrimination laws and additional regulation to curtail the use of potentially discriminatory algorithmic or AI evaluations. Public policy documents highlight the growing need to address the use of AI in high-risk contexts, such as education, employment, health, housing, law enforcement, loan underwriting, justice, and public benefits [71, 161, 296]. In these contexts, decisions or evaluations made through or with the assistance of AI systems could significantly impact people's economic well-being, health, legal circumstances, and educational opportunities. A recent representative survey of Americans found that most people would be uncomfortable with the use of AI or computer algorithms to make high-stakes decisions about their lives [138].

Our work focuses on education and employment decision-making contexts in which voice-specific AI applications could help evaluate human performance in hiring and college admissions interviews, employee performance evaluations based on presentations, and grading of students' pronunciation on English final exams. In these contexts, evaluation may be based on factors such as pronunciation [32, 87, 111] and communication skills [132, 396].

### 3.4.2 Sociotechnical Harms of Algorithmic Systems and AI

Prior work on algorithmic harms has conveyed how marginalized communities disproportionately experience sociotechnical harms resulting from algorithmic systems [123, 136] and that "algorithmic systems' enactment of power dynamics [155, 309] can function as a minoritizing practice [85] through which unjust social hierarchies are reinforced" [373]. Additionally, AI systems may not perform as well for certain groups of users and thus result in further marginalization of vulnerable communities. For example, an AI system might incorrectly classify a subject's race or gender due to inadequate training on similar subjects [137, 369], which may create different, and potentially unfair, outcomes for different populations.

These harms can span a wide range of contexts, including employment and hiring, as well as in other contexts like content moderation, as addressed in prior CSCW literature. For example, social media content creators have expressed concerns about algorithmic reinforcement of

marginalization based on their identity or content, through algorithmic content moderation that limits their reach to broader audiences and "shadowbanning" [92, 96, 150], potentially limiting or silencing the voices of creators in marginalized communities. Not only could this cause emotional distress or frustration, but also lead to economic damages should content be removed or excluded from platform monetization programs [203]. On the other hand, AI systems can perpetrate harmful beliefs by aggregating and reducing entire groups based on stereotypes, particularly those demonstrated on social media platforms [243]. While these specific contexts are not included in our study, algorithmic suppression of free expression on social media and the reduction of groups based on stereotypes may lead to similar harms as those surfaced in our work, including economic impacts from limiting monetized social media posts or societal-level effects from the perpetuation of stereotypes.

In our study focusing on high-stakes employment and education contexts, we categorize participant responses using some of the harms provided by Shelby et al.'s taxonomy of sociotechnical harms of algorithmic systems based on a scoping review of 172 research articles [373]. In these contexts, Quality of Service harms, such as service or benefit loss, may manifest in differing levels of performance for speakers of non-standard dialects, potentially leading to unfair evaluations in employment or education contexts (Section 3.6.4 (a)). Allocative Harms, such as opportunity loss or economic loss, could be a result of differences in performance, meaning that those who are unfairly evaluated may miss out on valuable job opportunities or be unfairly excluded from suitable colleges (Section 3.6.4 (b)). Interpersonal Harms, such as privacy violations, may occur if the data collected in sensitive contexts (e.g., interviews for jobs or college admissions) is leaked or improperly protected Section 3.6.4 (c). Representational Harms, such as stereotyping, and Societal Harms such as cultural harms of systemic erasure could occur as the result of unfair evaluations over time, potentially leading to the exclusion of certain speaker groups from employment or education opportunities (Section 3.6.4 (d)).

Our study specifically focuses on potential algorithmic harms based on dialect, accent and speech. While there is less research on language-specific algorithmic harms, some prior work has increasingly brought to light the need to consider the treatment of language and dialect in AI technologies. In a study on fairness datasets and what attributes protected by anti-discrimination legislation across multiple continents they include, Simson et al. show that "language" as a feature was available for only three of 36 fairness datasets and used in only one of 233 experiments [375]. Thus, language, and dialect, is an understudied characteristic in fairness literature.

### 3.4.3 Uneven ASR Performance for Certain Dialects and Accents

AI applications using automatic speech recognition (ASR) can help assess human standards of pronunciation based on statistical models of desired outputs [32, 87, 111]. However, prior work and reports have shown that ASR performs worse for non-standardized English dialects and accents [56, 208, 366, 404], languages that do not have a significant amount of existing data for models to be trained on ("low-resource languages") [340, 341], and non-standardized dialects and accents of non-English languages [121, 325]. Though previous work has provided insight into improving ASR by going beyond a traditional computer science approach [252, 253, 254], other work has indicated the potential difficulties or harms that could result from attempting to overcorrect training data gaps using questionable means [313, 314]. Additionally, accent itself

may be ill-defined, and Prinos et al. conducted a survey of ASR literature and found that accent is often not considered or described in-depth, with a lack of clarity on what is considered the "baseline" accent or the incorrect assumption that some users may not identify with any accent at all [325].

A common standard for measuring accuracy and performance is called word error rate (WER). A higher WER implies a worse performance of an ASR model. Voice assistants have been shown to have higher WER for speakers of certain dialects or accents, resulting in failure to understand users' commands [34, 251]. Prior work specifically on US English dialects has also shown disparate WER. Martin and Tang analyzed over 100 hours of spoken AAE using two ASR systems, focusing on a "unique morpho-syntactic feature of [AAE]," the habitual "be" and found that the ASR systems showed a higher error rate around instances of the habitual "be" than for the non-habitual "be" [257]. Koeneke et al. found that five off-the-shelf ASR systems show disparities in transcribing identical phrases uttered by Black speakers compared to White speakers in an experimental study with 42 White and 73 Black speakers, with a corpus spanning five U.S. cities [208], and they suggest that including AAE in training datasets could reduce these performance differences. Some work in this space seeks to improve model performance, including developing new forms of measuring model perfomance [398], and robustness for different dialects [110, 386, 387].

Prior work has also emphasized the need for dialect-sensitive and culturally aligned technology, with some researchers taking a human-centered approach to inform the development of dialect-sensitive technology. For example, Harrington et al. and Brewer et al. interviewed Black older adults in the US regarding their use of voice assistants and found that some participants expressed perceptions of a language disconnect, noting that they had to "code switch" for the Google voice assistant, from AAVE, or what one participant called "Black-sounding" speech, to SAE [45, 148, 149]. Markl et al. explored why commercial ASR systems and other language technologies perform worse for marginalised second-language speakers and speakers of stigmatized varieties of British English and the policy implications [251, 253] and propose "an interdisciplinary and context-sensitive approach to documenting [systemic predictive bias for marginalised speaker/user groups]" and argue that "evaluation of ASR systems should [...] consider user experience in a broader sociolinguistic and social context" [252]. Other NLP researchers have consulted speakers of non-standardized dialects to better understand their perceptions of failures [159] and their preferences for model alignment [359].

### 3.4.4 Underrepresented or Marginalized American English Dialects

While it is widely acknowledged that "the typical NLP [Natural Language Processing] pipeline underrepresents speakers of most of [the 7000 human languages] while amplifying the voices of speakers of other languages," it has also been shown that some dialects of high-resource languages are underrepresented in NLP pipelines. Wale et al. analyzed a "typical NLP pipeline" and found that "even when a language is technically supported, substantial caveats remain to prevent full participation" [425]. For example, prior work on toxic or hate speech detection has suggested that performance varies based on dialect [127, 364]. Thus, AI systems for high resource languages, such as English, may not adequately serve all speakers of those languages. Researchers have called for an interdisciplinary approach to natural language technologies, noting that improving

| Dialect Group | Criteria Based on Self-reported Presentation |
|---|---|
| African American English (AAE) | Speaks AAE often/always<br>• May also speak SAE or Southern often/always |
| Appalachian English (Appalachian) | Speaks Appalachian often/always<br>• May also speak SAE or Southern often/always |
| Southern English (Southern) | Speaks only Southern often/always<br>• Not AAE, Appalachian, or SAE |
| Standardized American English (SAE) | Speaks only SAE often/always<br>• Not AAE, Appalachian, or Southern |

Table 3.1: Participant groups based on self-reported dialects spoken: African American English (AAE), Appalachian English (Appalachian), Standardized American English (SAE), and Southern English (Southern).

AI models is not just an engineering problem [370].

We selected participants from the southern U.S. because we sought participants who might be familiar with non-standardized dialects, including Southern American English, Appalachian English, and African American English (AAE), all of which are linguistically associated with the American South [8, 13, 217, 218, 219, 336]. These dialects are sometimes stigmatized in comparison to the "Northern" or "standardized" American dialects, resulting in historical patterns of discrimination [202, 317, 327]. However, even within the Southern US, there are varieties of dialects and accents that may be considered more or less stigmatized or prestigious based on region [62, 83].

It is also important to acknowledge the diversity of language within social groups and not essentialize ethnicity as representative of any given dialect. Some critics highlight the overly generalizing aspects of grouping the speech of African Americans across the U.S. into one dialect, suggesting that this leaves out nuances based on region or other factors [349, 438]. Similar criticisms have been made for delineating languages in general [151, 376].

## 3.5 Methods

In this section we describe how we recruited participants, designed and conducted the survey, and analyzed the data.

### 3.5.1 Recruitment

We recruited participants using Prolific's screeners and our own paid screening survey. Eligibility criteria included being at least 18, being raised and currently residing in the southern U.S., and speaking American English. Specifically, the following states where southern U.S. dialects are spoken [219, 336] were chosen for recruitment: West Virginia (WV), Kentucky (KY), Alabama (AL), Arkansas (AR), Georgia (GA), Louisiana (LA), Mississippi (MS), North Carolina (NC), South Carolina (SC), Tennessee (TN), and Virginia (VA). A gender-balanced sample was obtained

using Prolific's "Quota Sampling." A second sample with the same criteria but restricted to those over 60 was drawn to better balance age in our participant pool.

Participants were paid $0.80 via Prolific to complete our 4-minute screening survey. This survey ensured eligibility criteria was met, checked understanding of what voice assistants (VAs) are, and collected self-reported dialectal background. We specifically asked participants if they used any of the following dialects and, if so, how often: African American English (AAE), Appalachian English (Appalachian), Southern American English (Southern), and Standardized American English (SAE). Participants who met our eligibility criteria, selected the correct definition of a VA, and self-reported speaking at least one of the four dialects "often or always" were selected for the main survey and placed into one of four dialectal groups, as described in Table 3.1. More information on recruitment is provided in Table B.3 in Appendix B.3, and screening survey questions are included in Appendix B.1.

### 3.5.2 Survey

Our survey was designed to gain insight into 1) acceptability for general and voice-specific applications of AI in education and employment decision-making contexts, voice data collection by certain entities and inferences made by voice assistants, as well as 2) anticipated benefits and harms of these applications. Participants answered Likert and free-response questions about their attitudes and potential benefits and harms of each use case of voice technologies. Participants were paid $3.00 via Prolific to take this 15-minute survey. All survey questions are included in Appendix B.1.

**Acceptability Attitudes (RQ1)** We asked participants to rate acceptability on a 5-point Likert scale from extremely acceptable to extremely unacceptable. We first asked them to rate the acceptability of data collection by the following groups related to AI-enabled decision-making in education and employment contexts: companies that build and sell voice assistants (e.g., Google, Apple, Amazon), companies developing AI models (e.g., Open AI, Google, Meta), school districts, potential employers, current employers, and teachers. We then asked about acceptability of the following potential inferences made by voice technologies about speaker features: age, dialect or accent, ethnicity or race, gender, health condition, mood or emotional state, region of origin (e.g., where participants grew up), sexual orientation, and speech or voice disorder.

We then asked participants to rate the acceptability of algorithmic decision-making software in high-stakes education and employment contexts, with two use cases per context, based on prior work (Section 3.4.1). The four use cases were presented in a randomized order to control for possible order effects. We avoided explicitly asking about AI to not prime them with a potentially controversial term. The four use cases are presented in Table 3.2 and Appendix B.1.

Each use case contained a general scenario in which "software" was used to evaluate data subjects, which participants rated on a 5-point Likert scale from extremely acceptable to extremely unacceptable. We then presented voice-specific software and described how speech data would be used in its evaluations. Participants were asked whether the inclusion of the voice-specific application made the scenario *less acceptable*, *more acceptable*, or if it made *no difference in acceptability*.

Following the acceptability questions for each use case, participants were presented with their general scenario answer choice and voice-specific answer choice and asked to "explain [their] overall attitude for this scenario and voice-specific application."

| Context | General Scenario | Voice-Specific Application |
|---|---|---|
| College Admissions Interview | A small liberal arts college uses software to determine eligibility for admission. | The software evaluates audio recordings of admissions interviews to determine suitability for small classes with interactive discussions. |
| English Final Exam | A high school district uses automatic grading software to grade an annual English final exam. | The software grades English word pronunciation in the spoken component of the exam. |
| Hiring Interview | A company uses software to rank the top job applicants throughout the job application process. | The software analyzes the job applicant's speech during a job interview to rate how understandable their speech will be to customers. |
| Job Performance Evaluation | A company uses software to evaluate current employees' job performance. | The software analyzes speech during presentations to evaluate public speaking and communication skills. |

Table 3.2: Four education and employment use cases of AI shown to all participants. Participants rated the acceptability of each general scenario and then rated whether they found the corresponding inclusion of a voice-specific application more or less acceptable for that use case.

**Potential Benefits and Harms (RQ2)** For each use case, participants were asked to list at least one potential harm and at least one potential benefit related to the given use case, (or, if they did not believe there were potential harms or benefits, to explain why).

**Speech-specific Attitudes and Experiences (RQ3)** As part of our instructions, we included a note, "We're especially interested in your attitudes, opinions, or experiences relating to the southern United States and dialects or accents found in this region." Through this prompt, we hoped to elicit participants' attitudes and experiences related to non-standard dialects, accents and speech.

### 3.5.3 Data Analysis

The first and second authors analyzed participants' accceptability explanations and potential benefits and harms using emergent/bottom-up coding methods. Themes that emerged from the survey responses were grouped as coding progressed. All responses were double-coded, and differences were resolved through discussion. Quantitative information is provided about the themes identified for descriptive purposes only.

We organized themes that arose in our analysis of participants' self-reported overall attitudes into four top-level categories of 1) negative opinion or critique, 2) positive opinion or support, 3) humans and technology, and 4) speech-related comments, each with specific subcodes developed throughout the coding process. Our code book with the overall attitudes, harms, and benefits codes can be found in Table B.1 in Appendix B.2.

### 3.5.4 Limitations

As other scholars have done [35], we consider the participants to be a non-probability sample, so their responses are not meant to generalize to broader population trends, but rather to provide insights into complex social phenomena. We report numbers for descriptive purposes to highlight differences in the participants' responses and do not claim that these results provide predictive power or could be used in probabilistic classifiers, for example, as the assessment of statistical regularities for this context is beyond the scope of this work. We have little reason to believe that the outcome variable of acceptability is operationally consistent across demographic factors, such as dialect.

Participant categories were defined based on the self-identification of participants as speakers of certain dialects, and we did not "validate" their dialects. We utilized dialect categories in this work to recruit a diverse participant pool based on self-identification, but we recognize that these categories are not precise and may not fully capture participants' dialectal characteristics. We also acknowledge that Southern English may be more accurately considered as a group or family of dialects that could overlap with some Appalachian or AAE dialects, as shown by participants who identified as speaking multiple dialects.

Additionally, our sample was relatively young, and most participants did not report having problems being understood by voice assistants or humans. Thus, our study did not capture many responses from older adults and people who report being misunderstood most of the time. We encourage future work that focuses on these specific groups to better understand their opinions on acceptability for voice data analysis in high-stakes use cases.

Our team of four researchers conducts research matters using SAE. Our members' dialectal background is as follows: [anonymized for submission].

## 3.6 Results

Below we summarize our survey results. We first provide participants' self-reported demographic information and experience with voice assistants (Section 3.6.1). We also summarize the results of the Likert scale responses about the acceptability of AI-enabled software in each use case, data collection by entities relevant to those use cases, and potential inferences about users of voice assistants (Section 3.6.2). We then summarize our qualitative analysis of participants' free-text responses stating their attitudes and anticipated benefits and harms across all four use cases, organized into broader thematic categories of: potential benefits (Section 3.6.3), potential harms (Section 3.6.4), speech or dialect specific concerns (Section 3.6.6), and criticism of technology's inadequacies that emphasize the need for human evaluation (Section 3.6.5). We present themes across all four use cases. When quoting participants, we specify the participant number and use case.

Counts of thematic codes across the four dialect groups were compared and did not indicate any stark patterns, but we provide some notable highlights in Figure 3.1 and Section 3.6.2 (c) and as follows: the potential benefits discussed by a higher ratio of self-reported SAE-speakers than any other group were: *adequate measure*, *efficiency*, and *improve skills*, and the potential harms were *errors or glitches*, and *unfair outcome*. The topics for which the SAE group ratio was the

lowest of all four groups were: *dialect, accent or speech* and *speech is not a proxy for intellect or skills.* See Appendix B.2 for definitions of these codes.

### 3.6.1 Participants

All 111 participants were American adults from the southern U.S. who speak American English as their first language. We provide summary proportions for dialect, U.S. states, ethnicity, gender, education, income, and prior use and attitudes toward technology in Appendix B.3.

Participants mostly identified as one of two ethnicities, with 56 identifying as White, 53 as Black or African American, one as Asian, and one as American Indian or Alaska Native. A total of 64 participants identified as female, 44 as male, three as non-binary, and two as transgender. Most participants were relatively young, with 62 participants reporting being up to 40 years old. Most participants reported their highest level of education to include at least some college education, with 52 reporting having obtained an Associate's or Bachelor's degree, 30 having attended college with no degree, 15 obtaining a high school degree, 11 obtaining a master's degree, two having less than a high school degree, and one obtaining a trade school degree. In terms of income, 56 participants reported that they made less than $50,000 per year, 51 between $50,000-$150,000, and three above $150,000.

Most participants had experience with voice assistants, with only three participants reporting having never used a voice assistant. Most participants reported using a voice assistant on their mobile phone ($n = 91$) or having encountered a phone-based voice assistant ($n = 93$). When asked if they had ever trained a VA on their voice to improve the VA's ability to recognize them, 60 participants responded Yes, and 44 responded No.

In free-text responses, 15 participants mentioned prior experience with or knowledge about AI applications similar to those in our study's four use cases. Some current AI applications they mentioned include algorithmic review of resumes or videos submitted by job applicants, automatic exam grading, and ranking of college applicants. P33 wrote,

> I've had companies demand I do extremely similar stuff, even talking about myself on camera without any sort of prompts or questions, only to have my applications denied within minutes and the video never even landing on the hiring admission's radar. (P33 Hiring)

Regarding job performance evaluations, P110 suggested that "using software to track work performance is pretty standard these days." In exams, P14 and P46 recalled negative experiences with their writing being "misjudged" by AI.

### 3.6.2 Acceptability Attitudes (RQ1)

We summarize our study's 111 participants' Likert scale responses to questions about the acceptability of the use of AI-enabled software in four decision-making use cases, data collectors, and inferences made by voice assistants (VAs).

#### 3.6.2 (a) Unspecified "Software" Acceptable but Analysis of Voice Data Less Acceptable

When asked about the general use of software in decision-making for hiring, job performance evaluation, and college admissions use cases, over half of participants found these uses to be *acceptable* (somewhat or extremely). For the final exam use case, just under half found this scenario *acceptable*. For AI-enabled analysis of voice data across all four use cases, less than a quarter of participants rated the voice-specific application *more acceptable*, with the rest being roughly split between *less acceptable* and *makes no difference*. These results are shown in Figures B.3 and B.4 in Appendix B.4.

#### 3.6.2 (b) Data Collectors: Companies More Acceptable Than Teachers, Schools or Employers

When asked to rate the acceptability of data collectors potentially involved in the four use cases, AI companies (n=66) and VA companies (n=58) were rated as acceptable by the most participants, followed by teachers (n=48) and current employers (n=44), school districts (n=31) and potential employers (n=29).

#### 3.6.2 (c) Some Inferences More Acceptable Than Others

When asked about the acceptability of inferences made by VAs, inferring dialect or accent was found *acceptable* by the most participants ($n = 76$), followed by gender ($n = 67$), mood or emotion ($n = 62$), age ($n = 61$), region of origin ($n = 59$) and speech disorder ($n = 58$). Less than half of participants found inferring ethnicity or race ($n = 43$), health conditions ($n = 25$) or sexual orientation ($n = 25$) *acceptable*.

For eight of nine questions about inference making, the SAE group had the least number of people and lowest percentage of their group who found the inference making unacceptable, as shown in Figure 3.1. However, it is important to note that these are small and non-representative samples of 32 or less people per group. with the following group sizes: Southern: 24, Standard: 25, Appalachian: 30, AAE: 32.

### 3.6.3 Anticipated Benefits: Efficiency, Objectivity, Improving Skills (RQ2)

Participants suggested that AI applications could increase efficiency ($n = 82$) by saving time and labor, increasing speed, or generally streamlining processes. This could relieve teachers, hiring managers, and administrators of tasks such as grading or reading applications. Participants ($n = 32$) also suggested a potential benefit of mitigating human biases by providing objective, fair, and consistent assessments or by not showing bias or favoritism. Additionally, participants suggested that AI applications could help employees or students improve skills ($n = 30$), such as providing feedback to students on their pronunciation. P38 suggested "it would be beneficial for students who are mispronouncing words. They can have it played back to them to see what they sound like" (Exam). Other participants noted how the software could broadly improve communication skills across both the employment and education contexts.

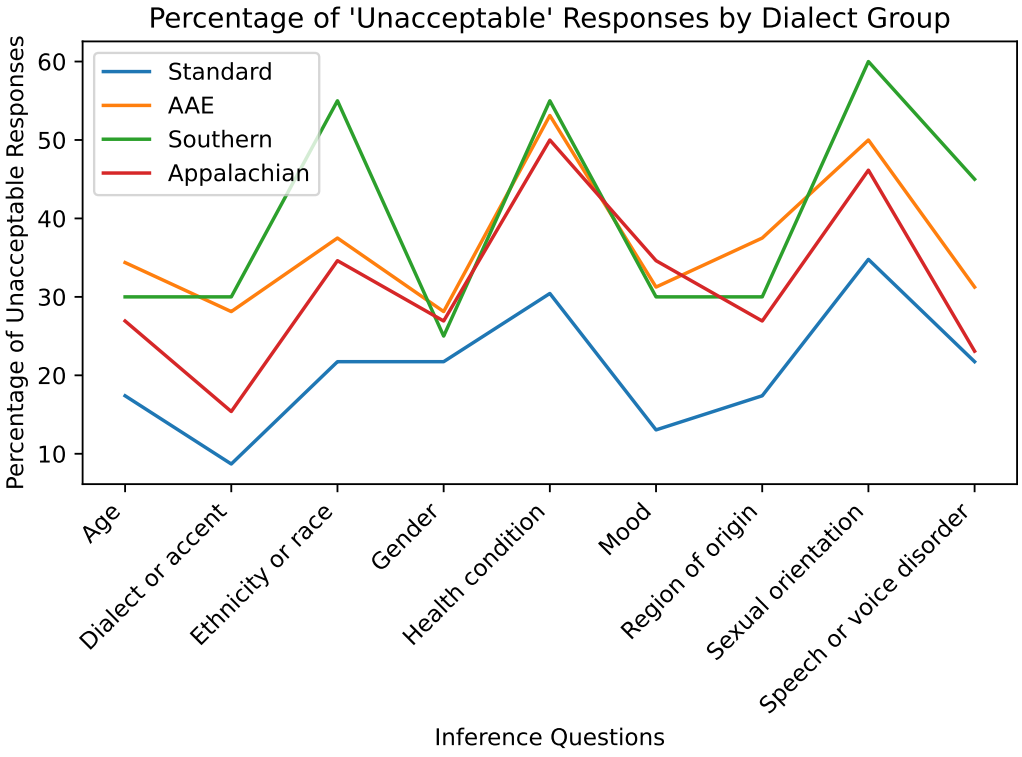


Figure 3.1: Line graph showing percentage of percentage of unacceptable responses by dialect group. The Standard or self-identified SAE group had the least number of people find it unacceptable to make inferences.

While not necessarily a benefit, we note that when discussing benefits, some participants expressed confidence in the ability of software, computers, and AI to evaluate understandability, communication skills, and pronunciation ($n = 43$) adequately or even better than a human could.

### 3.6.4 Anticipated Harms: Unfair Outcomes, Discriminatory Bias, Privacy Issues (RQ2)

A total of 90 participants mentioned potential unfair outcomes, which primarily consisted of unmerited or erroneous negative decisions for each use case: rejection of job or college applicants, lower final exam grades (and thus a lower overall grade), and negative job performance evaluations resulting in the denial of pay raises. Such potential outcomes were sometimes linked explicitly to discrimination but also included outcomes resulting from error, malicious intent, inadequate training of the AI technology, or generally "unfair" without specific cause.

Additionally, 79 participants specifically named the outcomes of discrimination or bias. Among these, across the four scenarios, 67 participants indicated a basis for discrimination, such as accent or disability, and 40 suggested there might be discrimination or bias without naming a basis. The use case that garnered the most comments about discrimination or bias was hiring ($n = 52$), followed by exam grading ($n = 49$), college admissions ($n = 47$), and job performance evalua-

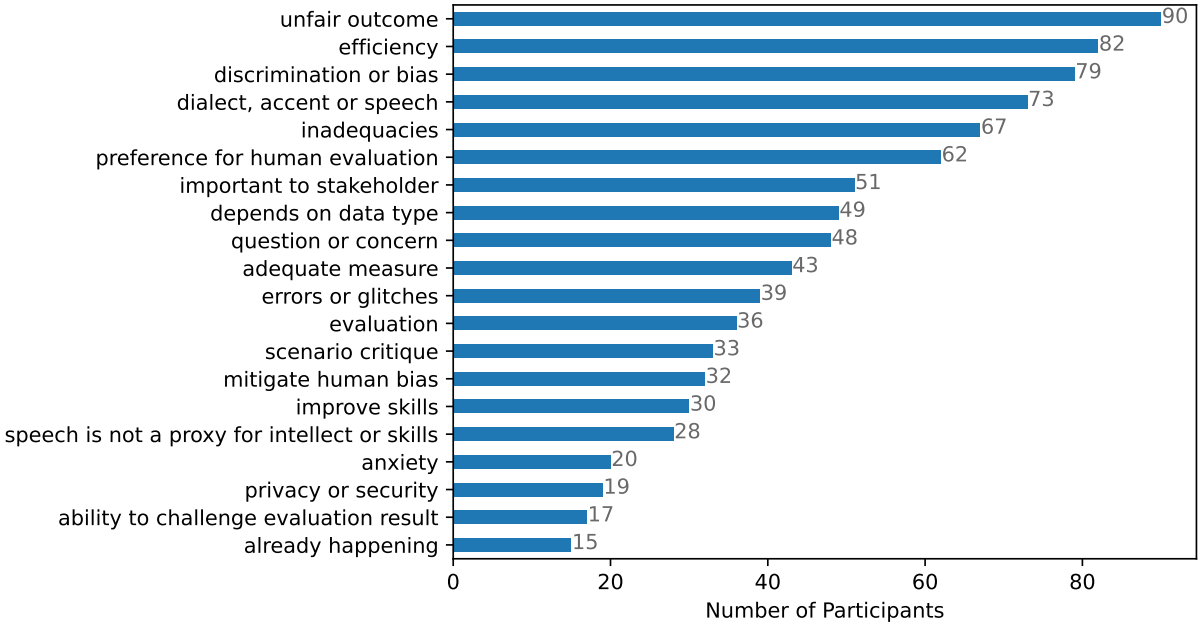


Figure 3.2: Count of themes, counted once per participant across all four use cases. See definitions in Appendix B.2.

tion ($n = 35$). Finally, 19 participants mentioned privacy and security harms. Below, we discuss quality of service, allocative, interpersonal, and societal harms raised by participants.

### 3.6.4 (a) Quality of Service Harms: Concerns of Worse Performance Based on Accent

Some participants discussed how AI-enabled voice data analysis could fail those with different accents, dialects, or speech disorders due to its inability to properly analyze them. For example, P34 wrote that a speech disorder or "heavy accent" should not negatively affect a student's grade, suggesting that such an application "may also disproportionately affect foreign students or students with different cultural background" (Exam). Furthermore, P59 expressed that "the voice assistant may hear a voice that they haven't yet heard, and then give that candidate a bad score. When in reality, that candidate is great for the role, but the voice assistant just hasn't heard that specific dialect or accent before" (P59 Hiring). This implies a lack of confidence that AI is capable of processing certain differences in speech.

### 3.6.4 (b) Allocative Harms: Negative Impacts on Educational and Job Opportunities

In the context of discrimination, participants mentioned concerns about disparate outcomes that could impact one's livelihood due to unfair rejections of qualified job or college applicants. For example, P34 expressed concerns about how this technology would "lead to less job opportunities for people who have accents or speech impediments" (Hiring). Another participant suggested that college applicants who speak non-standard regional dialects might not be considered "'articulate" enough (P55 Admissions). Overall, participants raised concerns suggesting that voice-sensitive

AI could lead to harms that restrict access to educational or financial resources, particularly for those who are already part of marginalized communities.

#### 3.6.4 (c) Interpersonal Harms: Privacy and Security Risks

Other individual-level harms mentioned by participants included privacy or security concerns related to data collection in each use case. In particular, participants considered how voice recordings would be handled or stored, malicious uses of their voice data like deepfakes, the potential for data breaches, and other privacy violations based on a lack of consent for the data to be used elsewhere. For example, P108 believed that voice recordings could be"used for other purposes later that employee did not authorize" (Job Evaluation). P27 shared similar concerns for the hiring use case: " I can only imagine some higher up having some nefarious plan on how to monetize the recorded interviews." Other respondents pointed out issues related to data storage, such as P22, who asked about the college admissions interview, "What if you [don't] get admitted? Where does your voice recording go?" P108 was concerned that recordings would be stored "too long, and [be] susceptible to a breach" (Hiring).

Some responses highlighted the risk of inference based on dialect. P95 suggested that exam grading of pronunciation "could result in the person's race or ethnicity being discovered, risking possible racism from the grading of the exam." Such inferences could be maliciously used by humans to make prejudicial decisions, potentially harming individuals further by restricting their opportunities (Section 3.6.4 (b)).

#### 3.6.4 (d) Societal Harms: Potential Perpetuation of Negative Stereotypes

Participants also recognized harms on the societal scale, such as perpetuating or creating harmful stereotypes. For example, some responses included suggestions that accent or dialect could be associated with negative perceptions about intelligence, articulateness, or acceptability. For example, P10 wrote, "i am worried that they may think i am not smart just because of my accent" (Admissions). P17 said she had a 3.8 GPA in high school but shared how various personal factors, such as her gender, "would be used to assume I have lower intelligence level automatically, especially my accent" (Admissions). Similarly, for job performance evaluations, P49 wrote, "People with unacceptable dialect could be discriminated against," suggesting that some dialects may be considered unacceptable in the workplace. Furthermore, a participant who self-identified as speaking Appalachian often or always listed discrimination and prejudice as potential harms of AI analysis of college admissions interviews, adding, "I was once told 'they were surprised I wore shoes"' (P72). These responses suggest that participants believe there is a potential for impacts on cultural groups by either continuing to impose existing stereotypes or perhaps even creating new ones.

### 3.6.5 Preference for Human Evaluation

In addition to participants' responses that address benefits and harms, we present a recurring theme found in participants' responses: a preference for human evaluation. Over half of participants (n=67) expressed skepticism about technology's ability to perform nuanced evaluations

needed for decision-making, as well as a preference for human evaluation (n=62) in the high-stakes decision-making use cases described in our survey.

Criticisms of technology's inadequacies included an inability to accurately interpret or quantify human emotions, body language, differences in speech due to dialects and accents, acceptable differences in speaking style and presentation, or other factors. P99 suggested that human reviewers "will be more aware of biases," adding, "[I]f say you were talking to someone of your own race, it may or may not be clear what you're saying to the computer, but a human will understand more" (Hiring). P55 also expressed doubt that an AI program would be able to understand "different ways of speaking" and that "it seems like a lot of minorities would be negatively impacted" (Exam).

AI's inadequacies also included the inability to understand charisma, personality, or potential for growth. For example, P38 wrote that technology used in hiring interviews "can't understand charisma. … The way one's voice makes people feel is important." P113 also indicated that, for the hiring use case, AI "may overlook vital intangible qualities like personality fit or potential for growth," implying that human reviewers could account for these qualities in their hiring decisions.

#### 3.6.5 (a) Inimicable Human Capabilities

Participants named human skills that would be hard to imitate and tasks that should be human prerogatives, such as where humans should have the upper hand. These skills included tolerance for differences, subjectivity, detecting nuance in pronunciation, reading body language or tone, having empathy for others, and taking additional factors into consideration. For example, P38 highlighted that "a foreign born student may technically not speak as well as a US one, but a teacher could give them the same grade taking it into account. It is relative to each student" (Exam). P6 suggested that AI could not sympathize or empathize, writing, "[I]t is a cold lifeless ai that cannot make moral judgments and can only go by some algorithm" (Hiring). P27 asked, "Why use software when managers can use the power of observation to determine how well an employee is doing[?]" (Job Performance). P33 criticized the absence of human review for applicants who submit videos for college applications, calling it "completely dehumanizing, because it shows the applicant that the admissions office can't even be bothered to speak with the applicants directly." These responses evoke the notion of dignity and a wish to be seen or heard by another person, highlighting the humanity of the act of giving each other time and consideration.

### 3.6.6 Attitudes and Experiences Related to Speech and Dialect (RQ3)

In addition to concerns relating to language or speech, such as discrimination based on differences in speech, participants also conveyed attitudes relating to speech evaluation in general. For example, about a quarter of participants (n=28) objected to the use of speech as a proxy for skills or intelligence. Some participants also provided insights into their background and experience as speakers of non-standard dialects, highlighting both negative and positive aspects.

### 3.6.6 (a) Resistance to Speech-Based Evaluations

Participants who objected to voice-specific evaluations to gauge skills or intellect suggested that speech does not represent the qualities most important to the evaluation. For example, P17 wrote that, for work presentations, "the actual presentation shouldn't matter, the results it presented as a result of the projects an employee worked on should be more important" (Job Performance). P82 wrote, "speech does not equal performance" (Job Performance). P72 expressed that they "have been made fun of for my accent my entire life. Yet, I have found that intelligence has nothing to do with accent. If an AI assistant judges based on accent, it needs to be scrapped" (Admissions). P19 suggested that "someone like [Stephen] Hawking," a renowned physicist and author who had a motor neuron disease that affected his speech, "would end up being declined by the software, although he's one of the smartest humans to ever live" (Admissions).

For work presentations, participants also raised potential factors of shyness, nervousness, register, or formality that might negatively influence a job performance evaluation. P45 wrote, "The practice is unfair to more reserved employees who are good at their job but lack good presentation skills" (Job Performance). P40 wrote, "Sometimes people are nervous when speaking in public but that doesn't have any bearing on their communication skills" (Job Performance). P19 suggested that work presentations would not be "a good representation of public speaking skills" because of the informal register they tend to use at work, writing that they "talk a lot less formal" at company meetings "since we all know each other" (Job Performance).

Some participants highlighted positive aspects of speaking non-standard dialects, such as P81's experience "that AI voice software has a problem with my southern/Appalachian dialect. But, I've worked in customer service for 20 years and have found customers can understand me and many say they love my accent" (Hiring). P71 noted the dialectal variety of Southern English, for which there is "a strong Scottish and Irish influence in some places and more of a drawl in others. These are fine and beautiful. If it's judging on the content of their speeches and not the twang, then it's okay" (Job Performance). They also suggested that voice-specific AI evaluations of hiring interviews "might be useful if you need people to take calls from a specific area of the United States," suggesting that such evaluations could promote regional dialects by inviting their usage in workplaces, as opposed to encouraging the use of SAE.

### 3.6.6 (b) Support for Evaluations Based on Standard Dialect

Some responses expressed support for evaluations based on the standardized dialect of SAE. For example, P53 wrote that voice-specific AI applications "could definitely help weed out employees with sloppy speech patterns, or ones that are too quiet to understand" (Hiring). P87 wrote that a benefit would be that "there will be people working for the company that speak clearly and are well understood by the larger population of people where this company is providing it's services" (Hiring). Addressing exam grading, P99 wrote, "I don't believe there will be any harm because in those scenarios you're supposed to talk proper even if you do have a slang or accent." Similarly, P68 wrote, "if you are using words, you should know how to pronounce them" (Exam). P25 wrote that potential college students' "[s]peech should be clear and concise to be admitted" (Admissions).

Some participants wrote about the potential for speakers of non-standard dialects to learn to

speak SAE. P71 suggested that the English exam grading program could help students "get better at faking their accent so that in interviews and phone calls, judgmental people don't automatically assume folks from this area are uneducated" (Exam). Furthermore, P60 suggested that college applicants with accents would benefit from the use of this software because it "would encourage such students who have deep accents to add another dialect to their portfolio, 'proper English,' which I admit is understandable to almost anyone who know[s] English" (Admissions).

However, P72 highlighted the difficulty for some speakers to adopt a more standard dialect and suggested that this inability to adapt linguistically might have repercussions on their job performance evaluations, writing that "the way we speak" would influence the outcome of the evaluation, continuing, "I have spent years trying not to sound like I was from southern [West Virginia]. Yet, I still do."

## 3.7 Discussion

Below we discuss takeaways from our results, including societal implications of AI-enabled analysis of voice data for for language ideologies and potential future interdisciplinary work related to dialects, how our contributions align with existing algorithmic harm taxonomies while also providing rich, context-specific data, and policy implications.

### 3.7.1 Risk of Reinforcing Standard Language Ideologies

Responses expressing support for evaluations based on a standard dialect (Section 3.6.6) emphasize "proper" and "clear" English, assume widespread acceptability of SAE, and depict non-standard speech as potentially "sloppy," poorly enunciated, or stigmatized. Participants' responses evoked existing power dynamics, social stigmas, prejudices, and negative personal experiences related to speech. AI-enabled technologies used to evaluate metrics such as understandability or pronunciation utilize statistical models trained on data that may consider a standard dialect as "correct" pronunciation, raising the possibility of representational harms, which "occur when algorithmic systems reinforce the subordination of social groups along the lines of identity" [373].

Participants also raised concerns about the potential systematization of dominant language standards, which could result in erasure of their dialects, due to a need for data subjects to conform to speech standards. Additional concerns arose regarding the potential for these systems to reinforce negative stigma associated with certain dialects or regions, relating to existing work on algorithmic harms of systemic erasure [97, 396] and "proliferating false ideas about cultural groups" [102, 362].

### 3.7.2 Future Participatory Work that Celebrates Dialects and Audits AI Tools

We encourage future work that emphasizes the celebration and use of minoritized dialects and accents to validate and promote them in ways supported by speakers, such as work that addresses gaps in NLP performance for "low-resource" languages [3, 187] or regional accents and dialects

[7, 9, 188], the preservation of languages and dialects [321, 412, 429], and harnessing voice technology to better understand and correspond with minoritized dialect groups [126, 322].

In our study, some participants wrote about positive aspects of using non-standard dialects, such as regional specificity and charm. In the spirit of *unmaking* literature and work that emphasizes joy, assets, and "the everyday" in the lives of marginalized people rather than only focusing on harms [37, 108, 405], we propose work that unmakes standard language ideologies embedded in language technologies [439] and engages participants using participatory, ethnographic, and archival [184, 287] methods to examine the perspective of self-identifying members of sociolinguistic communities. Such work could consider language usage, local history, personal narratives, archives and other information to inform the design, or rejection, of language technologies [453].

We also encourage the development of user-centered auditing of language models and language-sensitive AI systems for speakers of low-resource or marginalized languages or dialects. Prior work on user-centered evaluation of language models [97, 232, 441] has emphasized the need for human-centered research and design that engages stakeholders to bring awareness to features, opportunities, and concerns that previously may not have been considered. Cultural and linguistic knowledge can help evaluate and red-team AI systems, as linguistic variations may cause a model to fail, throw errors, or produce unfair outcomes. Such variations could include phonetic differences, morphology specific to a given dialect, or lexical semantic variations. Given that some of our participants reported experiencing difficulties being understood by voice assistants or having faced discrimination, our findings suggest that auditing with linguistically diverse populations could help make AI and ASR systems more robust and fair.

### 3.7.3 Contextualizing Algorithmic Harm Frameworks

We identified examples from sociotechnical algorithmic harm categories defined by Shelby et al. (Section 3.4) in our data, which allowed us to identify how quality-of-service harms where systems perform differently for certain groups (Section 3.6.4 (a)) can cascade into other harms, such as unfair evaluations (Section 3.6.4 (b)) and perpetuation of negative sterotypes (Section 3.6.4 (d)). We provide contextualized considerations and details grounded in real-world problems and experiences. For example, these education and employment contexts, algorithmic harms could not only damage a person's professional or academic career but also result in legal troubles for institutions deploying such technologies.

We also contribute themes outside of a general harm taxonomy. For instance, many participants indicated a preference for continuing human involvement, with or without the presence of AI, across all use cases (Section 3.6.5), highlighting how relying on AI-enabled analysis may not be adequate measure for certain goals. While creating more inclusive voice technologies can help ensure evaluations are more fairly conducted, participants pointed out how this kind of AI evaluation does not account for factors such as potential for growth. Some participants objected to voice data serving as a proxy for skills (Section 3.6.6).

This desire for continuing human involvement suggests an inimitable benefit or contribution provided by humans that allows for non-data-driven qualities to be considered. A lack of human involvement, where evaluators such as teachers or managers cannot easily address or correct biased results in proprietary AI models, and where they cannot apply their human discern-

ment, could further harm those impacted by disparate or discriminatory outcomes, echoing prior work [18]. We echo a call in recent work to and investigate how decision makers and recruiters approach AI-assisted evaluations, including how they "use and make sense of AI systems, and how this affects their discretionary decision making" [377, 378] and to "better understand how hiring managers, workers, applicants, and others within an organization interact with each other and with hiring tools" [379].

### 3.7.4 Policy Implications of Voice-specific AI Applications

Concerns raised by participants about the risks of augmenting or automating employment decisions with AI have implications for equal opportunity and antidiscrimination policy. Indeed, in the last few years, various countries and regional governments have passed regulations addressing this problem [345]. In 2024, the European Union adopted the the AI Act, providing guidance on high-risk applications of AI systems [82]. In the U.S., cities such as New York City and the states of Utah and Tennessee have passed laws regarding automated employment decision-making tools [434]. In Illinois, the Illinois Human Rights Act was extended to include protections against the AI-enabled discrimination based on racial data or common proxies, such as ZIP codes, in several employment use cases, including those related to hiring and job evaluation processes [2, 120]. In Colorado, the Colorado Artificial Intelligence Act (CAIA) provides standards of use and disclosure when AI is involved in high-stakes contexts, such as employment and education, to prevent discrimination of any kind [348, 389]. This law includes procedures required both for creators and users of this technology, which would include employers or schools relying on AI for hiring candidates or assigning major grades [348, 389]. The California Civil Rights Council has advanced regulations regarding algorithmic decision-making systems in the workplace [294, 319], which have since been passed [1]. We support the further development of such policies to help safeguard people from AI-based discriminatory outcomes.

Participants also raised privacy and security concerns regarding the collection and storage of voice data, and potential malicious use or breaches of this data. Similar to our results, prior work found that potential users of biometric inference systems considered AI-based biometric inference-making to be privacy invasive [15, 213, 356]. Some scholars have suggested that machine learning is inciting a new age of phrenology, with biometric data being used to draw conclusions about people, sometimes erroneously [141, 225]. We recommend that researchers and companies question and investigate potentially false claims correlating biometric data with ambiguous variables [147, 379], such as personality and hireability [346, 347]. Additionally, given the potential integration of ASR into AI systems, we call for policies that codify protections for voice data as biometric data. Biometric data regulation has the potential to curtail the irresponsible or high-risk utilization of voice data [358]. Given the potential integration of ASR into AI systems, we also encourage future work that helps protect voice data through anonymization and other privacy-preserving features for voice technologies [6, 115, 130].

## 3.8 Conclusion

Our survey study considers the attitudes and opinions of people from the southern U.S. regarding the collection and use of their voice data by AI-enabled technologies in four high-stakes decision-making use cases. In considering four scenarios related to education and employment, participants anticipated benefits that could increase efficiency and objectivity. However, harms that could reinforce and systematize bias, result in discrimination, and pose privacy and security risks were also discussed. Thus, we recommend the development of AI applications by and for speakers of non-standard dialects, user-centered AI auditing by these speakers, and increased legal protections.

# Chapter 4

# Emerging Technology Context: IT/OT Integration in Critical Infrastructure

This chapter was adapted from my published paper:

> Andrea Gallardo, Robert Erbes, Katya Le Blanc, Lujo Bauer, and Lorrie Cranor. 2024. "Interdisciplinary Approaches to Cybervulnerability Impact Assessment for Energy Critical Infrastructure." Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, Article No. 828, 1–24.
> https://doi.org/10.1145/3613904.3642493

## 4.1 Overview

We first present a study consulting stakeholders about security problems posed by the convergence of information technology (IT) and operational technology (OT) in energy infrastructure. This IT/OT convergence is an increasingly significant security problem with potential large-scale implications, as the energy grid becomes interconnected with emerging technologies such as cloud systems and AI tools.

This work connects to the thesis as an example of engaging stakeholders whose lived experiences and expertise allow them to articulate context-specific security challenges and opportunities. In individual interviews with two types of subject matter experts—computer security researchers and energy operational technology professionals—we discussed the problem facing many stakeholders in critical infrastructure: how to assess the potential impact of IT vulnerabilities in OT systems. The two groups' different backgrounds and yet overlapping interdisciplinarity provided rich insights into challenges and possible solutions in securing energy OT systems. Using qualitative analysis methods of a priori coding and emergent bottom-up coding, we analyzed participants' stated approaches to vulnerability impact assessment and their perceptions of each expert group, documenting similarities and differences across the groups.

Our findings highlight both similarities and differences in their impact assessment approaches and the way they discuss security. Participants, all of whom had interdisciplinary experience

working with the other kind of expert, displayed cross-domain knowledge and raised security topics at similar rates, suggesting that their overlapping interdisciplinarity provided them with cross-domain knowledge relevant to securing energy OT systems. Yet, we also observed notable differences in how these two groups of experts discussed their approaches, using discourse analysis and emergent coding to reveal gaps in understanding, training, and occupational priorities. For example, cyber SMEs displayed a more adversarial focus on gaining access and modifying device capabilities, and energy OT SMEs expressed more holistic considerations regarding the impact on the overall system and potential disruptions to operations.

Our findings provide empirical data on gaps in understanding and mental models of experts whose work directly impacts the security of the energy grid. Further work building upon this study could inform communication, education, and coordination efforts that enable disparate groups of experts to collaborate effectively to secure critical infrastructure. Given that protecting energy systems will require utilizing knowledge from both groups, we also recommend future work analyzing differences in communication, epistemology, and culture among these groups and developing interventions to bridge the cross-domain knowledge gap.

Our final study builds on this study, in which we interviewed hard-to-reach stakeholders by establishing trusted relationships with contacts who helped recruit participants, and also analyzed culturally situated perspectives and language.

## 4.2 Abstract

As energy infrastructure becomes more interconnected, understanding cybersecurity risks to production systems requires integrating operational and computer security knowledge. We interviewed 18 experts working in the field of energy critical infrastructure to compare what information they find necessary to assess the impact of computer vulnerabilities on energy operational technology. These experts came from two groups: 1) computer security experts and 2) energy sector operations experts. We find that both groups responded similarly for general categories of information and displayed knowledge about both domains, perhaps due to their interdisciplinary work at the same organization. Yet, we found notable differences in the details of their responses and in their stated perceptions of each group's approaches to impact assessment. Their suggestions for collaboration across domains highlighted how these two groups can work together to help each other secure the energy grid. Our findings inform the development of interdisciplinary security approaches in critical-infrastructure contexts.

## 4.3 Introduction

Knowledge sharing and collaboration between energy operators and computer security professionals is needed to understand risks to and potential impacts on energy production systems. The protection of energy infrastructure is an immensely critical computer security problem. Disrupting energy grid operations can have particularly severe consequences for society, with loss of power potentially causing a ripple effect that impacts other critical sectors and services, such as hospitals [278, 350, 433, 443], financial services [223, 344, 406, 448], agriculture [84, 99, 256,

352, 372], and energy production and distribution [106, 241, 305, 388].

However, while these two groups of experts need each other in order to secure energy systems, they come from different disciplines, operational cultures, and sometimes have competing motivations and approaches (e.g., block connections vs. keep connections open for remote maintenance, patch immediately vs. schedule downtime to patch). In energy operational contexts, the security of electric-grid equipment has often been considered in terms of equipment failure or misuse, as energy systems were traditionally independent of information technology (IT) or relied on barring connections to external networks [48, 273]. IT security approaches and frameworks are often inadequate for energy operational contexts, which face challenges such as legacy systems that run on old operating systems and the need to operate continuously, which can delay patching and updates. Additionally, the security of energy-grid operational technology requires an understanding of how this technology is responsible for the generation, transmission, and distribution of energy and how computer vulnerabilities can impact these energy-production processes.

Nevertheless, the operational technology (OT) used in such critical infrastructure increasingly relies upon computers and computer networks to operate, as systems like power grids become integrated with networked Internet of Things devices and require maintenance through connected devices or remote workers. Thus, energy OT infrastructure becomes increasingly vulnerable to attacks through exploitation of computer vulnerabilities [388].

However, there is a well-documented shortage of computer security professionals [36, 171, 227, 233, 308, 343], and smaller energy facilities and utilities may lack resilient defenses and recovery plans [426] due to limited economic, staff and computer security resources [186]. Finding ways to build cross-domain knowledge will allow low-resourced utilities to help their staff make better-informed decisions about how to address risks posed by computer vulnerabilities, and also help computer security experts, whether on-site or designing industry-wide standards, develop security measures that are suitable for energy environments. While it may not be reasonable to expect OT engineers to perform the roles of computer security professionals or vice versa, building each group's awareness of risk factors in the other group's domain could help them seek appropriate resources to address risks to the energy grid.

Given this disciplinary divide between energy operational engineering and computer security and the need to develop cross-domain considerations, we situate our work around 18 employees of an energy-sector organization from these two domains, to compare their approaches to assessing the impact of computer vulnerabilities on energy OT. By computer vulnerability (hereon, vulnerability), we mean an exploitable weakness in a computer system, system security procedures, internal controls, or implementations that could be exploited or triggered by a threat source [303]. More specifically, these subject matter experts (SMEs) were: 1) computer security experts who primarily perform research in industrial control system security (cyber SMEs) and 2) operational technology experts with experience in engineering and operation of energy systems (energy OT SMEs). Our research questions are as follows:

- RQ1: What information do cyber SMEs and energy OT SMEs need when assessing the potential impact of computer vulnerabilities? Are there notable differences between the groups (i.e., cyber SMEs and energy OT SMEs)?
- RQ2: What do these experts consider to be the differences between the two groups' approaches to impact assessment and understanding of vulnerabilities?

- RQ3: What insights or suggestions do these experts provide that directly address collaboration between the two groups or building cross-domain understanding?

When self-reporting their approaches to impact assessment, both groups responded similarly at a general level, with roughly the same number of experts per group discussing each vulnerability impact assessment topic we coded. Both groups displayed knowledge about both domains, perhaps due to their interdisciplinary work at the same organization.

Nevertheless, we observed notable differences in the details of their self-reported considerations, as well as in their perceptions and suggestions regarding both groups' impact assessment approaches. These differences regarding each group's domain-specific focus and understanding were particularly interesting given that all participants had cross-domain work experience. Differences that shone through despite interdisciplinary backgrounds, such as cyber SMEs' more adversarial focus on gaining access and modifying device capabilities or energy OT SMEs' holistic considerations about connections across the system and potential disruptions in operations, highlight some domain-specific aspects that could be harnessed in complementary ways for critical infrastructure security. Indeed, many participants emphasized the value of cross-domain dialogue and exposure to the other group and had several suggestions for collaboration, building mutual understanding, and improving usability and security in energy OT contexts.

Our findings inform design for interdisciplinary security in critical infrastructure contexts by characterizing experts' approaches to impact assessment and highlighting differences in focus, mindset and understanding. Echoing suggestions made by participants, we recommend bringing experts together to foster cross-domain exchanges, developing training, tools, and educational interventions to help interdisciplinary practitioners build cross-domain understanding, and implementing usable security solutions in energy OT contexts.

## 4.4 Related Work

Our work provides insight into key issues in interdisciplinary impact assessment in energy OT contexts, focusing on the differences in approaches, professional motivations and skills of two groups of experts: computer security researchers whose work primarily focuses on vulnerability analysis and energy operational technology engineers. Below we discuss prior work establishing differences between OT and IT security, including perceptions and biases. We also note some existing frameworks and prior work on assessing risk or impact in computer security and energy OT contexts, as well as studies regarding cyber SMEs' and non-experts' mental models and perspectives in computer security contexts.

### 4.4.1 Contrasting OT and IT Security

Prior work has shown there are major differences between security approaches in IT and OT contexts, including differences in workers' training, knowledge, and culture, regulations for IT security versus OT safety, and conflicts between IT policies and OT continual operations. Studies have considered differences or conflicts between security approaches to IT and OT systems [73, 94, 158], as well as differences between considerations for operational safety and computer security in critical infrastructure OT systems [157, 240, 273, 437].

Wolf et al. identify key differences between traditional IT security and physical industrial control systems (ICS) computer security problems and make recommendations for remediation during design and runtime. For example, they discuss the potential for false data injection to cause harm to physical systems by creating unsafe operational conditions despite not traditionally being considered by cyber security threat models [437].

Prior work has noted historical and cultural differences between OT engineers and IT workers, suggesting that mindset, training, and epistemological approaches differ considerably [273, 282, 337]. Studies on collaboration and communication in security contexts have established a disconnect between IT security professionals and non-security professionals [339]. Michalec et al. highlight the historical differences between security incidents in IT and OT systems, given that "these systems were traditionally built for different purposes," and argue that there are "epistemic and material differences between legacy OT environments and big data practices." In their study interviewing 30 critical infrastructure OT professionals, they show that security risk management practices in critical infrastructure, which often relies on "old world" legacy systems, "cannot be directly transplanted from the safety realm, as cyber security is grounded in anticipation of the future adversarial behaviours rather than the history of equipment failure rates" [273]. The authors highlight three collaborative aspects critical to risk management across security and safety: access to diverse expertise and professional practices, trust and engagement between IT and OT workers, and the collective development of "risk thinking hiveminds," i.e., sharing expertise and best risk management practices across the sector via working groups.

While prior work has considered differences between IT and OT workers' security management practices, e.g., what mitigations are acceptable, how often to patch, and who has responsibility, our work considers how risk or impact is understood and assessed by experts working with OT systems. Rather than focusing on IT security professionals whose job responsibilities may include setting organizational IT security policies or monitoring networks for anomalous behavior, we consider cyber SMEs whose work identifying and analyzing vulnerabilities is distinct from energy operations and yet increasingly necessary to prevent, mitigate, and resolve computer security problems in energy OT equipment and systems.

### 4.4.2 Risk or Impact Assessment

The most commonly used framework for assessing the severity of vulnerabilities is the Common Vulnerability Scoring System (CVSS) [298, 304, 440]. Hollerer et al. attempted to merge CVSS and two safety and security frameworks to develop a "risk evaluation methodology to prioritize and manage identified threats considering security, safety, and their interdependencies" [157]. Prior work has also looked at ranking vulnerabilities in critical infrastructure [20, 89]. Some research has considered impacts or risks of interdependencies in critical infrastructure systems [284, 409]. Prior research has also provided suggestions for how to determine cyber security risk for energy sector infrastructure [11, 41, 170, 178, 221, 452] and OT systems, such as supervisory control and data acquisition (SCADA) systems [60, 86, 157, 333, 390]. While there are industry-wide reliability standards, such as the North American Electric Reliability Corporation Critical Infrastructure Protection (NERC CIP) standards [291], there is no industry standard for assessing risk for OT systems that is as widely adopted as CVSS is for scoring the severity of vulnerabilities.

### 4.4.3 Subject Matter Experts

Our study is concerned with the opinions of experts and how those opinions can be used to inform the vulnerability impact assessment process. Prior research has studied the mental models and skills of cyber SMEs and their perspectives on certain problems [10, 26, 27, 113, 177, 250, 397, 408, 423, 424, 436], suggesting that the perspectives of cyber SMEs can be valuable in outlining issues for certain tasks and that knowledge required for understanding vulnerabilities can often be specialized and varied. For example, in 2018, Votipka et al. showed that there is a divide between how two domain experts, "testers" and "hackers," think about software vulnerability discovery, and explored differences in factors such as training and motivation. They found that "hackers" were better able to spot vulnerabilities than testers [423]. Botta et al. and Hawkey et al. interviewed IT security professionals to characterize their responsibilities, goals, tasks, and skills, as well as difficulties collaborating within organizations [42, 154]. Reinfelder et al. interviewed seven IT security managers and found that IT security managers had difficulty receiving adequate feedback regarding usability of security features [339].

Comparative studies between experts and non-experts have also helped identify gaps in security approaches [39, 177]. Prior work has also considered the computer security perspectives of other kinds of subject matter experts, such as network administrators [210], Internet Service Providers (ISPs) [385], data scientists and data engineers [279], and cryptographic library developers [182].

Some qualitative studies have provided insight into practitioner perspective on critical infrastructure computer security. Line et al. assessed preparedness through interviews about situation awareness and incident response [239]. Michalec et al. interviewed 30 cyber security practitioners for their views on directives standardizing computer security for critical infrastructure [274]. Reilly et al. conducted interviews with 31 relevant stakeholders, including critical infrastructure operators, regarding how crisis information is communicated in critical infrastructure settings [338]. Yet, as far as we know, prior work does not provide detailed insight into the perceptions, experiences, and suggestions of energy OT SMEs and cyber SMEs regarding cross-domain collaborations assessing the potential impact of vulnerabilities on energy OT systems.

## 4.5 Methods

Below we describe our participant selection and recruitment process, our interview protocol, how we analyzed data, and limitations of our study.

### 4.5.1 Participant Selection

Study participants consisted of two kinds of experts: power systems experts ("energy OT SMEs") and computer security experts ("cyber SMEs"). The energySMEs were mostly engineers who maintain or manage energy systems. The cyberSMEs were researchers who utilize their deep understanding of how vulnerabilities work, harnessing skills like reverse engineering, to discover and analyze vulnerabilities in devices or systems they assess on a workbench.

We recruited participants from lists developed by colleagues (who were SMEs themselves)

at Idaho National Laboratory, a U.S. Department of Energy national laboratory that conducts research on energy and national security. Each list consisted of people who qualified as one of the two types of experts based on their current professional responsibilities and department, i.e., their current work took place primarily in one of the two fields of expertise. We did not share the list of suggested potential participants beyond the two authors who conducted interviews. Most participants responded directly to a recruitment email, and a few responded to follow-up emails from these two authors, who were not managers and did not work directly with participants. Managers were not involved in the recruitment process to avoid any sense of coercion. Participants were informed in recruitment materials and the consent form that the study was voluntary, and they had several opportunities to decline to participate or request that their data be deleted. They participated during work hours, and their employer paid them their normal salaried rate for the time they spent on the study. Only the two authors who conducted interviews had access to the deanonymized videos and transcripts. The only data and findings shared with the employer were anonymized results. All study protocols were approved by both the Carnegie Mellon University and Idaho National Laboratory institutional review boards.

### 4.5.2 Interviews

We conducted semi-structured interviews to capture the nuanced thought processes of experts as they considered their approaches to vulnerability impact assessment. Each interview lasted between 60 and 90 minutes and took place via Microsoft Teams between November 2021 through April 2022. All but one interview were recorded, and all interviews were automatically transcribed by Microsoft Teams software (including the non-recorded one). Transcripts were subsequently reviewed and corrected by the first author, based on recordings and notes. Our interview questions are included as Appendix C.3.

We began each interview by collecting general information, such as occupational background, years of experience, and experience conducting impact assessments. To better protect the identities of participants, we did not collect gender, age, income, or education level, though to the best of our knowledge, every participant had at least a bachelor's degree.

We then elicited and discussed the individual SME's general strategies for assessing the impact of a cyber vulnerability, what information they would need, and how subsector, context, vendor, and other factors might influence their approach. We also asked questions to elicit the SME's perceptions of differences between the two SME groups, i.e., differences in approaches to assessing the impact of vulnerabilites and differences in understanding of vulnerabilities.

### 4.5.3 Data Analysis

We structured our analysis around strategies, perceptions, and suggestions. We developed two codebooks for this analysis: one contained a list of *a priori* codes for impact assessment strategy topics that was refined throughout the coding process. The other codebook contained codes that emerged from review of the transcripts.

| Code | Description | Example |
|---|---|---|
| Accessibility | Information on the reachability of the vulnerable system. | Can I talk to the system from the internet? Is there an attack vector that can reach the system? |
| Attack | Understanding of adversarial threat, consideration of attacker, attacker motive, or actions. | How appealing is the system to an attacker? Who is the attacker? |
| Connectivity | Information on what the system is connected to. | Is the system connected to more important systems? What does the vulnerable system talk to? |
| Consequence | The result or possible result of malicious action upon the vulnerable device. | How long will it take to recover from an attack? Who would an outage affect? At what cost? |
| Consult other SME | Seeking external expertise outside of the participant's domain. | I would need to ask a power engineer to understand what would happen. |
| Device Information | Information about the system or device the vulnerability was identified within. | What does the system do? Where is it typically used? How common is it? |
| Vendor | Information on or about the company that builds the vulnerable system (unprompted). | Does the vendor provide support? What is their track record for fixing vulnerabilities? |
| Vulnerability | Information about the vulnerability itself. | Severity rating (e.g., CVSS). Can it be exploited? |

Table 4.1: Definitions and examples of top-level strategy codes. Subcode definitions can be found in Appendix C.4.

#### 4.5.3 (a) Impact Assessment Strategy Topics

The first three authors developed an initial list of codes based on their technical and research experience in computer security and vulnerability analysis. These codes were intended to help us categorize and track the participants' stated approaches to vulnerability impact assessment by honing in on whether they discussed particular topics. For example, did they mention understanding of vulnerabilities, potential consequences like loss of power, or how the system in question connected to or controlled other things?

We then coded a few of the responses for self-reported approaches to assessing the impact of vulnerabilities, iteratively returning to the codes to discuss disagreements, refine or consolidate the codes, and add any codes we felt captured concepts not covered by the initial list. This helped us further develop main codes and subcodes. We used the code book developed in this process

to code all responses to the open-ended strategy questions as well as the questions that elicited participants' perceptions of their own expert group and of the other expert group. Our final list of codes for strategies or approaches to assessing the impact of vulnerabilities are described in Table 4.1. We also developed subcodes to capture more details about each category, described in Appendix C.4.

For each question, the first or second author assigned codes to the responses (coder), and the other author reviewed the codes, noting any disagreements or adding new codes (reviewer). We tallied final code counts separately for self-reported responses and perception responses. Each sub-code was counted only once per participant, even if it was mentioned repeatedly, to allow for clearer group comparisons. We also coded each response with a perception valence of positive or negative, where the term "positive" means that the described group would consider the factor, or would be effective at considering the factor, and "negative" to signify the converse. Since the two groups had consistent positive and negative views of their own and the other groups (see Appendix C.2), we report results for positive and negative perceptions in aggregate.

#### 4.5.3 (b) Group Perceptions and Suggestions

The first author, either as a coder or reviewer, also used the following codes to characterize participants' responses to the perception questions, developing codes in a bottom-up coding process by first identifying detailed themes and subsequently reviewing the responses to thematically group them into three categories: 1) Stereotype: the SME group tends to do certain things or see things a certain way; general characterizations. 2) Occupational Motivation: habits, mindset or approaches based on training or job; what they are expected to do. 3) Suggestion: a recommendation regarding interdisciplinary work or collaboration. Another researcher reviewed these codes to verify their appropriateness and to suggest changes or additional codes. All codes are included in Appendix C.4.

### 4.5.4 Limitations

Our team is composed of computer security researchers and one human factors researcher. One limitation of our work is that our development of thematic codes was informed primarily by a computer security perspective. There may be additional codes that could have been included, had an energy OT SME been on the research team.

Another limitation is our small sample size of experts, which limits the generalizability of the results. While we sometimes report counts to make it easier to understand whether opinions were unique or more widely held, we don't imply any further quantitative characterization of the responses.

Additionally, all participants came from the same organization and may share overlapping or similar interdisciplinary experiences that could inform their responses and thus diminish notable differences for each group. Finally, questions about their own and other SMEs' understanding and abilities may have lead to responses with social desirability bias.

## 4.6 Results

We first provide background information about participants' professional and interdisciplinary experience (Section 4.6.1). We then present results for participants' self-reported impact assessment approaches (Section 4.6.2) and results for responses to questions about their perceptions of SME groups' strategies and understanding (Section 4.6.3). Finally, we relay their suggestions directly addressing interdisciplinary collaboration in energy OT security contexts in Section 4.6.4.

### 4.6.1 Participants

We interviewed 18 participants including nine cyber SMEs and nine energy OT SMEs from the same organization. We provide background information by participant number in Table 4.2 and additional details below about prior experience in impact assessment and cross-domain experience. When we provide numbers of participants in parentheses to characterize the responses, we use "Cyber" to indicate cyber SMEs and "OT" to indicate energy OT SMEs.

Four participants had 1–5 years of experience, five had 11–15 years of experience, and five had 16–20 years of experience. The remaining four had over 20 years of experience. Seven participants (1 OT, 6 Cyber) joined the organization directly after finishing higher education, and 11 (8 OT, 3 Cyber) had work experience prior to joining the current organization.

#### 4.6.1 (a) Impact Assessment Experience

Nine participants (7 OT, 2 Cyber) had prior experience conducting impact assessments, and when asked if their work "focused on the impact of cyber vulnerabilities," four additional participants (all Cyber) said yes, and one energy OT SME with impact assessment experience said no.

When asked about standard impact assessment procedures, participants noted that there was no standard impact assessment procedure for energy OT environments, but they mentioned some standard tools that could be used for impact assessment: the CVSS scoring system [298, 304] (3 Cyber, 1 OT), Common Vulnerabilities and Exposures (CVE) [77, 301] or Common Weakness Enumeration (CWE) [78] reports (2 Cyber), the NIST Cybersecurity Framework [302] (1 Cyber, 1 OT), the MITRE ATT&CK framework [79] (1 Cyber), a methodology developed at the organization (2 OT), as well as North American Electric Reliability Corporation (NERC) Critical Infrastructure Protection (CIP) standards [291], US-CERT and ICS-CERT (now CISA) alerts and advisories [67, 68], and the CARVER methodology [38, 295] (each 1 OT).

#### 4.6.1 (b) Cross-domain (Computer Security and Energy OT) Experience

All participants had some cross-domain experience. Of the energy OT SMEs, all nine had on-the-job exposure to computer security, and seven had some exposure to vulnerability analysis. Of the cyber SMEs, all nine had on-the-job exposure to energy OT systems. All 18 participants had worked on the same team as the other kind of SME and had also either worked on the same project or did work that overlapped with the other SME group's work, requiring coordination or complementary approaches. Such an interdisciplinary group is not typical in energy OT contexts. We

thus want to emphasize that our reporting of differences and similarities is not meant to generalize to trends in the energy OT industry. Rather, our results provide details about responses from experts in a particularly interdisciplinary group. We hope their responses will provide insight valuable for future work on developing cross-domain knowledge both among interdisciplinary experts and in environments where working together is less common.

### 4.6.2 Self-Reported Impact Assessment Strategies (RQ1)

We report how participants responded to questions about what information they would need to assess the impact of a vulnerability in an energy OT system, highlighting similarities (Section 4.6.2 (a)), differences (Sections 4.6.2 (b)–4.6.2 (c)), and interdisciplinary knowledge (Section 4.6.2 (d)). Despite bringing up similar high-level topics, there were notable differences in participants' stated approaches to vulnerability impact assessment, indicated by the level of detail participants provided, such as how cyber SMEs had more specific considerations about gaining access to networks, or how energy OT SMEs spoke more about connections to the overall system and potential disruption of operations. While we include numbers in some of the results below and in Table C.5 in Appendix C.5 to characterize this particular participant pool, we note again that these are not generalizable results.

#### 4.6.2 (a) Similarities in Self-Reported Impact Assessment Strategies

We expected cyber SMEs and energy OT SMEs to show a stark imbalance in their stated approaches to vulnerability impact assessment, based on prior work (Section 4.4.1), but we did not find this to be the case. Experts across both groups raised similar vulnerability impact assessment topics relating to Accessibility, Connectivity, Consequence, Device Information, and Vulnerability (described in Table 4.1), as shown in Figure 4.1 and Appendix C.5. Responses to questions about whether subsector and vendor would influence their approach were also similar. We hypothesize this may have been due to all participants having interdisciplinary experience. We describe these similarities below.

**Accessibility** Both groups of participants spoke generally about how to gain access (remotely or physically), who might have access, and access controls.

**Consequence** Both groups were aware of possibilities for large-scale impact, emphasized understanding systemic and broader scale implications, and offered general considerations about the potential impact on human life, damage to equipment, financial or business impact, remediation or recovery time, and whether it would affect critical systems.

**Device Information** Participants from both groups said they would need information about what the system or device is, what its function is or how it is typically used, how it is configured, who uses it, where it is located and situated within the OT system, how widely it is deployed in the overall environment and in the country, and how well it is protected from a physical and network standpoint. One participant, E7, called for a software bill of materials (SBOM) to better

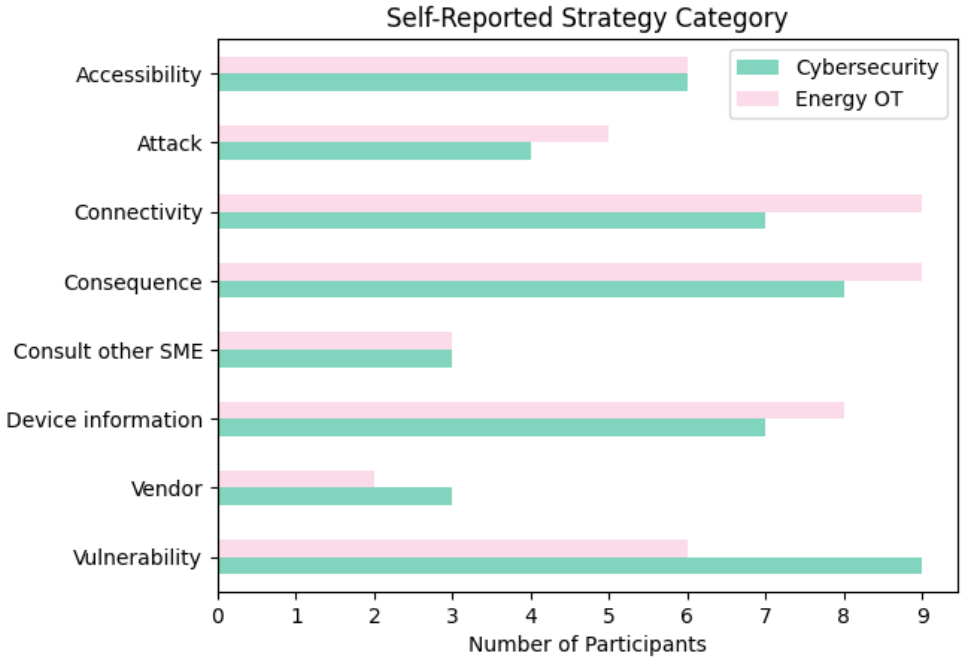


Figure 4.1: Top-level impact assessment topics considered relevant by participants in their self-reported impact assessment approaches, showing count of unique participants by SME group. We did not observe a stark difference between the two groups, which may have been due to the interdisciplinary background and experience of all participants.

understand where a vulnerable component exists and what the "most granular indivisible element that this vulnerability impacts" is.

**Vulnerability** Participants from both groups said they would consider what the vulnerability was, the area or systems affected by the vulnerability, proof of concept, and exploitability.

**Subsector** In response to questions prompting participants to tell us whether the energy subsector (e.g., generation, transmission, or distribution) would influence their vulnerability impact assessment approach, all but one participant felt that considering subsector was important for impact assessment, with one cyber SME saying the subsector would not impact their assessment at all (C3). In all cases where subsector was described as being important, it was due to the scale of the potential impact. Some participants ranked subsectors by importance, named the one they thought was most critical, or stated that all subsectors were equally important.

**Vendor** Five participants (2 OT, 3 Cyber) listed vendor as a factor they would consider before being prompted to discuss how a vendor would influence their vulnerability impact assessment approach. In response to our question, four participants (2 OT, 2 Cyber) stated that the vendor doesn't matter at all. All other participants stated that the vendor did matter. Some participants said different vendors had different track records with computer security, specifically with how

open and communicative each vendor was with their customers regarding vulnerabilities, emphasizing the importance of strong lines of communication and relationships. Two participants discussed the difficulties of trying to report vulnerabilities to different vendors (E10, C12).

### 4.6.2 (b) Differences in Cyber SMEs' Self-Reported Impact Assessment Strategies

Cyber SMEs' responses were distinguished by a more adversarial focus on gaining access, identifying connections, and imagining device capabilities and exploits.

**Gaining Access** Cyber SMEs spoke in more detail about gaining access to a system's networks or devices than energy OT SMEs, saying they would want information about an attacker's ability to move around within an OT environment (C13) or a company's networks, including SCADA or control system networks (C12), or the potential for an attacker to access "additional resources that you can chain to get there or tie in with other controllers" (C11).

**Identifying Connections** Cyber SMEs also emphasized identifying connections across networks and systems, imagining paths to hard-to-reach devices:

> *I would try to trace a path to this piece of equipment to try to understand how easy it is to get there. Some equipment is designed to be on a network that is more likely to have malicious traffic. Other equipment is not designed for that, and it's expected that it's going to be behind several firewalls. (C12)*

C11 said, "Often, end devices are not very reachable. So you'd have to have access to several other networks or a different entry point to get to them." Others asked questions such as "Does it cross the boundaries between different networks?" (C5) and "Can you go to the next system over?" (C15).

**Device Capabilities** Cyber SMEs provided more examples about what affected devices or systems might be capable of doing. Three cyber SMEs suggested devices could be modified to perform unintended actions or be mis-programmed by abusing a given functionality. C5 wanted to know potential impact "if it was to be modified and if the vulnerability allows the code to be changed and just run something arbitrary instead." C5 also considered the possibility of modifying a device or system that is "just supposed to be gathering data" to "send commands to something" on the same network. C12 considered looking for potential capabilities of hardware components:

> *Sometimes we think about a device as a single device, but if you open it up under the hood there might be two or three or four different devices inside with distinct functionality. And maybe one portion of that device is built to be more trusting, and so you have to look and do a divide and conquer approach. (C12)*

C12 also contrasted the idea of looking for a "novel exploit, which you should definitely search for," with an unintended or "insecure implementation" of a provided functionality. C17 considered the potential impact of "issues with the device," such as an "accessible debug shell" that could be made to "continually crash and restart, taking up resources."

In contrast, three energy OT SMEs spoke more generally about understanding what affected devices or systems were capable of doing, with E16 also considering the ability to change things "within the product to be used inappropriately."

**Exploit details** Cyber SMEs also considered details regarding potential exploits. C5 wanted to know if there were "creative" ways to modify the device or change the system's code and if this could take the system offline. C9 said they would consider what kind of data could be released by the vulnerability, as well as its severity. C11 said they would be more concerned if it were possible to "chain" the vulnerability with knowledge about other vulnerabilities to create a larger impact. C12 expressed concern about older systems being exploited with published "off-the-shelf" vulnerabilities. C14 asked whether the vulnerability was persistent or temporary and whether it could spread to other things.

All five cyber SMEs who raised the topic of exploitability asked how "easy" it would be to exploit the vulnerability, while energy OT SMEs asked whether it was "actually" exploitable (E7, E16) or would require remote access (E10). The cyber SMEs' responses implied that compromise was possible but that their consideration depended on difficulty, highlighting factors like how reachable the system is and the attacker's skill level.

#### 4.6.2 (c) Differences in Energy OT SMEs' Responses.

Overall, energy OT SMEs conveyed a more holistic view of the system and provided more concrete examples of parts of systems and how systems might be affected.

**Connections to the larger system** Energy OT SMEs provided more details about how connections between devices and systems relate to the overall system. For example, E1 considered an engineering workstation as "something with a pretty low impact for safety or operations" but with high potential security impact because "it touches everything" and might contain credentials and configuration files. E7 and E15 were concerned about whether the vulnerability was "on something centralized that controls a lot of different things, like my SCADA or EMS" (E7). E8 expressed concerns about distribution systems "becoming more integrated," saying, "Historically, a distribution system was one radial feed. Now it's starting to talk to all the meters out in these residential areas." E16 was concerned about effects on "the downstream load."

In contrast, cyber SMEs asked general questions such as what it means for devices connected to the system (C17), "what equipment is being used and what ties they have to the outside world or to any type of network" (C3), what the system communicates with, controls, or monitors (C5), and what the dependencies on the system are (C14), not providing potential answers themselves, as some energy OT SMEs did, implying that they would obtain this information from another source, such as an energy OT SME.

**Disrupting operations** Energy OT SMEs also spoke in more detail about potential disruptions in operations. For example, E15 considered whether the location might be a "high priority site" that needs to "maintain critical loads" and whether it would thus be among the last users to lose service and the first users returned to service after an interruption. E16 wanted to understand how

much power or the "amount of megawatts and gigawatts" that might be "turned off" and what point in power distribution was disrupted: a meter at a residence or a transmission substation.

Energy OT SMEs also wanted to understand what kind of disruption might occur. For example, E18 wanted to distinguish between whether the vulnerability would "completely shut us down" or "only slow us down temporarily." E7 suggested that temporarily mitigating a threat by "physically remov[ing] some kind of communication channel" might cause people to complain that they "need the data," but suggested that the potential impact might not be great: "But do you really need it? Are you billing from it? Is it a regulatory thing, or is it just something that you'd like to have?" Thus, for E7, potential impact on business processes might be higher impact than lacking nice-to-have information.

**Risk mitigation** When discussing vulnerabilities, energy OT SMEs also focused more on stopping or mitigating the vulnerability, containing the risk, patching, ensuring operational integrity, and understanding the residual impact or risk of the vulnerability for the larger system and operations. For example, E15 said that after stopping an attack, they would verify "the integrity of operational functions" and if they had control of all equipment and operational status, and then "find what was potentially accessible to the attack" and confirm protection systems were still functional and working "as designed" and "that my rules haven't been changed on my communication devices."

### 4.6.2 (d) Cross-domain Knowledge

Some participants displayed cross-domain awareness in their responses or said they would seek out such knowledge as part of their impact assessment strategies. C17 made a distinction between whether a system could be accessed by customers at their homes or by engineers at a generation plant:

> *If there's an exposed port that you can connect to that gives you debug access or a shell, that would largely be an issue with a consumer device, because that means your consumer could do whatever the heck they want to with your device. But in the case of a high reliability system in generation, it might be significantly more important to have that as a means of debugging any issues that do occur with the device. (C17)*

C12 said they would not consider a cabinet containing "a bunch of ethernet ports that you could connect to," to be "very high impact" if it were inside "a facility with 10 layers of physical security." Additionally, E6 considered the cyber hygiene of portable media and mobile devices accessing the system: "What do you do for maintenance? Do you bring a laptop over? Do you sanitize all of your portable media?"

Additionally, C14 and E7 conveyed cross-domain knowledge when speaking about isolating systems containing the vulnerability. E7 mentioned the "occasional" situation in which they are able to "wall off" a vulnerability "that's not actually used for the functionality of that product":

> *It is incredibly difficult and maybe in a few cases straight up impossible to actually exploit, and then that lets me back off and step back from the ledge a little bit and say "OK, this is important," but it's not like, "Oh my God," the end of the world here. (E7)*

C14 evoked concepts from a recent training on safety risks:

> *Going through lab training the other day, there's a whole, when you have a safety risk or safety issue, the best thing to do is to eliminate it. The second thing to do is to have controls that contain it. The third thing to do is to tell people not to use it. … So I guess that applies also in this kind of system. (C14)*

This suggests that C14 was applying knowledge from an energy OT safety training to a computer security context.

Additionally, one energy OT SME and one cyber SME participant suggested methods for obtaining interdisciplinary insight, emphasizing how they would consult the other expert group once they had seen proof of concept for the given vulnerability. E16 said they would consult a "product SME" and cyber SME to collaboratively understand what the vulnerability was capable of doing. C17 said they would "lean on the energy SMEs" to gain insight into potential implications for the system, how easy it would be to replace the device, if the device could be "ruined" by the vulnerability, or what kind of impact it might have "in terms of environmental impacts or larger societal impacts."

### 4.6.3 Perceptions of SME Groups (RQ2)

We first present recurring generalizations or perceived tendencies about each group, in Sections 4.6.3 (a)–4.6.3 (b) Because we didn't spot any particular difference between the characterizations advanced by the two groups of SMEs, we present the stereotypes by the group who is the target of the stereotypes. Then, in Section 4.6.3 (c), we convey what participants said were the occupational motivations, or driving factors, of the expert groups. Participants' positive or negative perceptions of both expert groups' vulnerability impact assessment strategies are included in Appendix C.1.

#### 4.6.3 (a) Stereotypes of Cyber SMEs

Some stereotypes about cyber SMEs were that they understand vulnerabilities, misunderstand energy OT systems and impact, reduce systems to computers, pay attention to details, overestimate impact, and cut off access to protect systems. Cyber SMEs were also seen as representing "IT" people or departments.

**Cyber SMEs' understanding of vulnerabilities and systems** Cyber SMEs were characterized as understanding exploits and vulnerabilities (3 OT, 4 Cyber), e.g., being able to tear devices apart to do things like extract firmware or find vulnerabilities. Ten participants said that cyber SMEs lacked understanding of energy OT systems (4 OT, 6 Cyber). Eleven participants said that cyber SMEs lacked understanding of impact or overestimated impact (5 OT, 6 Cyber), while only one said they underestimate impact (1 Cyber). C12 suggested that cyber SMEs "are more likely to think the sky is falling when it's not." E7 also suggested they may incorrectly think a vulnerability could crash the grid:

> *The cyber security folks tend to think of it as: "This is exploitable, and if you can do this, you can crash the grid with it." Whereas the electric folks are like, "OK, no. You can maybe knock off that one generator, but in reality, you can just knock*

*off the controller for that induced draft fan, and that means I would have to de-rate my generator. ... I'm not making as much money that day. But it's not the end of the world. (E7)*

Thus, this overestimation could be due to not understanding redundancies in place and safeguards that prevent a vulnerability from impacting systems.

**Cyber SMEs see computers.** Cyber SMEs were depicted as treating OT systems as computer systems that can be manipulated as such (1 OT, 3 Cyber). E7 conveyed how systems perceived by engineers in terms of their function could be reduced to modifiable computers. "From the perspective of the maker, the people who install it, [and] the protection and controls people," a protective relay is a device that quickly and reliably "reads electrical voltage and current," then "does some math on them" to determine whether or not "to send a trip signal to a breaker." Yet, they added:

*From the adversary, cyber security perspective, this thing is a computer. It's got a full-blown operating system. It's running Yellowstone Linux or Windows 8.1 embedded or something else. And if I have the right passwords or I can figure out how to bypass the different protections on it, I can make this thing do anything that a computer could do. (E7)*

**Cyber SMEs focus on details.** Four participants emphasized cyber SMEs' attention to detail (1 OT, 3 Cyber). C11 and C12 said they go into "rabbit holes" and that this could be both a good and bad thing, with C11 suggesting the importance of "reigning yourself in" when focusing too much on one type of analysis, and C12 acknowledging that some things may be interesting from a cybersecurity standpoint but may end up being low risk. Yet, they said, it is not always clear whether it is low or high risk until it is fully tracked. E1 said cyber SMEs spend months on device vulnerability analysis doing a full evaluation of a device. C13 suggested that cyber SMEs underestimates impact because they focus on "the here and now" details about the immediate environment rather than thinking about implications and how something might "cascade" through a system.

**Cyber SMEs cut off access to protect systems** Five participants said that cyber SMEs cut off access to protect system (3 OT, 2 Cyber). Several responses indicated that cyber SMEs were perceived as restricting access to systems in order to protect them. Indeed, some participants provided anecdotes or made suggestions that evoked frustration with a lack of usable solutions.

*There needs to be open communication between certain applications, certain devices. And completely locking those down, to the level that a lot of cyber security experts would like to see, just isn't feasible. ...A lot of times the OT, I think, just kicks out cyber security and says "Get out of my hair." (E8)*

**Cyber equals IT** Four participants (3 OT, 1 Cyber) discussed cybersecurity and IT departments in the same statements, suggesting an association between the two. When responding to a question about cyber SMEs, E6 suggested that "IT people" don't understand how controllers work

and how they communicate, and thus they "think that they can just go onto the OT side and do the same thing and then they have they find out the hard way":

> *They don't have the understanding of how controllers work and how they communicate, so if you run certain things to do analysis, you could potentially take out your production system, where on an IT system, it wouldn't matter (E6)*

C14 suggested working with IT teams to develop authentication solutions, since if "the IT Department is enforcing things without actually talking to the people who have to use it, then you never figure out that you can come up with different authentication systems." E15 also evoked IT being an enforcer when, for example, "IT says we've got to make a firewall or system" to avoid public access to the grid.

**Other Perceptions** Less common perceptions included that cyber SMEs overemphasize the following: complicated exploits when simpler ones achieve same effect (1 OT), IP-level communications (as opposed to serial and proprietary level communications) (1 Cyber), patching (2 OT, 1 Cyber), and software (1 Cyber). Two energy OT SMEs said that cyber SMEs underestimate the importance of continuous operations and keeping things functioning (1 Cyber) and underestimate or fail to consider misuse of technology (2 OT). One participant suggested that cyber SMEs lack funding or resources (1 OT).

### 4.6.3 (b) Energy SME Stereotypes

Energy OT SMEs were represented as understanding systems and impact, not understanding vulnerabilities or exploits, lacking imagination, and taking shortcuts.

**Energy OT SMEs understand systems.** A repeated opinion was that energy OT SMEs understand the design, maintenance, and operation of energy systems, energy OT equipment and capabilities, and how system components are connected. E15 also said energy OT SMEs know how to install equipment, maintain it correctly by making sure it integrates well with other equipment that's coming in, and replace old equipment. C11 highlighted how energy OT SMEs' input about systems helps them understand impact:

> *Usually the people that are talking about it and introducing us to it are quite honest about, "And if this part goes down, it's gonna be a huge pain." They may not be thinking about it in risk, but they usually point out parts that are difficult to replace or computer systems that are very key to keeping up and running. (C11)*

E4 also spoke specifically about asset owner operator energy OT SMEs, saying that they "will know their systems better than anyone else on the planet."

**Energy OT SMEs do not understand vulnerabilities.** Energy OT SMEs were characterized as lacking understanding of exploit capabilities, attacks, vulnerabilities, technical details, and network communications. Regarding overestimating and underestimating, seven participants suggested that energy OT SMEs underestimate the ease with which vulnerabilities could be exploited (1 OT, 6 Cyber). C5 suggested that people who set up the OT systems may not think about how easy it is for the different systems to be compromised and not consider lateral movement across

different parts of the network and creative ways of gaining access. Relatedly, three participants said that energy OT SMEs overestimate protections (1 OT, 2 Cyber).

Other things that participants said energy OT SMEs underestimate include: access & connectivity (2 OT, 2 Cyber), hardware attacks (1 Cyber), impact (2 OT), misuse (1 OT), and risk (1 Cyber). Participants also suggested that energy OT SMEs overemphasize the following: network security (1 OT), physical security (1 OT), prior vulnerabilities (2 Cyber), system resilience (1 OT), vulnerability score (1 OT), what a device is supposed to do (1 Cyber), and software (1 Cyber).

**Energy OT SMEs lack imagination.** Eight participants suggested that energy OT SMEs find it difficult to think of possibilities outside of what they already know (2 OT, 6 Cyber), i.e., are not able to imagine vulnerabilities or potential exploits or harms beyond what they already know.

> *Some of that stuff is not readily clear just by say reading about something or walking through its configuration, which is typically what an OT SME might do, where they don't go to that layer below, they really just look at what's there and kind of accept that that's how it is. (C13)*

E7 suggested that energy OT SMEs fail to see computers in OT devices: "It's not a protective relay. That's a computer. It can do computer stuff." C11 suggested that some energy OT SMEs needed convincing or explanations when told that a device had to be replaced due to a severe vulnerability.

> *Some OT SMEs do not believe you, they're like, 'Oh no, you just reboot it. It's made to be reliable.' ...They have worked with systems for a long time, and they are used to things breaking and being good after a reboot or two. They think that's the same thing for someone actively trying to exploit or damage a system. ...They're so used to it just being able to recover because it's made to have high reliability for the types of things that happen accidentally, that having someone purposely damage it is a completely foreign idea. (C11)*

Additionally, E16 suggested that energy OT SMEs might not consider "the potential downstream" impact of a highly motivated and resourceful attacker exploiting a relatively minor vulnerability on a large scale, e.g., rather than simply opening one switch, creating a scenario with "hundreds of or thousands of switches that are opened and then you can't re-close them because communication lines have been taken out." C17 suggested that energy OT SMEs see some systems as always failing into a known state or behaving in known and proven ways, and that they do not consider vulnerabilities that can change how the system behaves. E18 suggested that how specific responsibilities are distributed amongst personnel in an organization could influence energy OT SMEs' understanding of vulnerabilities: "You may have one person that is responsible for aspects of the operations on a daily basis. They need to make sure that the facility is functioning and might have less of a concern about quote unquote potential risk."

**Energy OT SMEs take shortcuts.** Six participants warned about energy OT SMEs taking shortcuts, thus leaving vulnerable defaults open, in order to work around restrictive policies put in place by IT or cybersecurity departments (3 OT, 3 Cyber). E1 said that engineers circumvent or work around security policies if they're too restrictive so that they can access the system, for

example, by putting in a back door. C14 said "at some point you're a little bit looser on your security because you need to get stuff done and there's other defenses."

**Other perceptions** Energy OT SMEs were also depicted as lacking funding or resources (2 OT, 1 Cyber) and "mistak[ing] safety systems being certified safe or a security certification with securing a system from hacking" (1 Cyber).

#### 4.6.3 (c) Occupational Motivations

Regarding impressions about cyber SMEs' occupational mission, ten participants suggested that cyber SMEs identify exploits, vulnerabilities, and flaws (4 OT, 6 Cyber). Three participants said that cyber SMEs tear apart or dissect systems (1 OT, 2 Cyber), and three said they focus on protecting computer systems (3 OT).

The most common impressions about energy OT SMEs' motivation were that they focus on making sure the system works and ensuring power delivery (5 OT, 3 Cyber). One or two participants also suggested that energy OT SMEs focus on the following: operational efficiencies such as maximizing reliability, minimizing costs, and saving time (2 OT), development or design (2 OT), protecting systems (1 OT, 1 Cyber), and human safety (1 OT).

### 4.6.4 Participants' Suggestions (RQ3)

Throughout our interviews, participants shared insights about their collaborative experiences working with other type of experts and made many suggestions for how collaboration, usability, design, and education could be improved. Indeed, when discussing their own strategies for impact assessment, six participants suggested they would consult a SME from the other group themselves (3 OT, 3 Cyber), and when discussing perceptions of group strategies, eight said that they expected SMEs to consult the other group (4 OT, 4 Cyber) in certain situations.

Overall, eleven participants made suggestions about collaboration and communication (6 OT, 5 Cyber), emphasizing bringing together the two SME groups to work on the same team or collaborate in shared work settings, opening up communication and listening to each other, and understanding the goals and areas of the other SME group.

#### 4.6.4 (a) Integrate OT Environments to Include Both Experts

Participants suggested integrating energy OT operational environments by having conversations that build mutual understanding, creating overlap in operational teams, and conducting red-team simulated attack exercises.

E15 suggested that one cross-domain problem is that the two groups have different conceptions of what it means to protect a grid: "I can talk to you about protection and line current differentials …, but to a cyber security person, it's not going to make any sense. But that's how I protect my grid. And they can talk to me in other terms about how they can protect the thing, that I'm not going to understand." E7 suggested that a way to build mutual understanding is to have a discussion between the two groups that "resembles a lot of the same processes that an adversary

would need to go through to develop a targeted capability to create a specific impact" and which requires both sides to go back and forth:

> *They're going to converge towards a point of understanding in ... that back and forth of, "What do you mean somebody exploiting this couldn't crash the grid?" Well, because that piece of equipment, no matter what you do with it, can't cause an electrical cascading event. "OK, I don't know what that means, but it makes me glad that you can't crash the whole grid. What can you do?" Well, here's all the things you can do with this equipment, if you had total control over it. And the cyber guy is like, "Here's what I would need to do to have total control over it." (E7)*

Thus, approaching issues "from different sides of the center" allows the interdisciplinary team to iteratively build understanding of potential impacts on the system.

E1 highlighted the importance of overlap in operational teams to securing critical infrastructure and said that "operational groups" should work with "the security side" in an integrated matter to understand risk and avoid "breakdowns" in operational contexts:

> *Why do I care if a cyber researcher understands power equipment, and why do I care if a power SME understands cyber? The only time I really care about that is implementing it in the actual operating utilities. If [those groups] can work in a more integrated manner, then they are going to do a better job. And engineers, if they understand the risks and the hazards, are very good at using that in their designs. But if they don't understand that risk, then they're going to exclude that. And that's what I think happens in a lot of places, that you don't have enough overlap, so even if your power SME wants to do things correctly, they don't understand how to do it correctly. (E1)*

E1 thus emphasized integrating energy OT operational teams to include cyber SMEs to help "utilities protect their systems operationally."

C12 said that in their experience, the situations where "knowledge tends to go back and forth" were exercises that brought the groups together and assigned them roles to attack and protect the system.

> *You have a group of people that will be focused on trying to break something, and then you would have a group of your OT experts that are there to manage the system and restore and reign in the other team from going too far and damaging everything. (C12)*

C12 said that through these exercises, "the cyber people would become more knowledgeable about the system, and the [energy] OT SMEs would become more knowledgeable about the cyber aspect and what could break."

#### 4.6.4 (b) Consult Other SMEs for Domain-Specific Knowledge

Participants emphasized the importance of consulting and listening to people who understand the other domain very well. Some cyber SMEs suggested that talking to energy OT SMEs who understand specific systems would help them understand how systems are set up or implemented, how they are supposed to work, how they actually work, and what they are connected to. C11 said that energy OT SMEs can identify which systems are key to keep up and running and thus

difficult to replace, as well as provide advice for how to replace such systems when they are no longer operational or should be put out of service due to vulnerabilities.

E8 suggested that they gained knowledge from cyber SMEs and said they now assume an air gap is already compromised: "I'm more cautious [about air gaps], but that's because I've worked with cyber security researchers and cybersecurity experts that told me otherwise. But that's not the commonly accepted practice, from my experience."

#### 4.6.4 (c) Make OT Systems More Usable

Seven participants suggested making energy OT systems more usable (5 OT, 2 Cyber). As discussed above, some participants conveyed the impressions that energy OT SMEs take shortcuts or leave vulnerable defaults open and that cyber SMEs cut off access to protect systems. E1 said that cyber SMEs "should not have it be so locked down that the operational side can't do their job," emphasizing that energy OT SMEs needs easy access to systems to do their job. C5 specifically spoke about usability, making the following suggestion:

> *If something is extremely difficult to do, but it has to be done all the time, then people are going to try to find shortcuts or ways around it, so [try] to work with the people who are using something to design solutions that will work for them. (C5)*

E8 said that there needs to be more open communication and more collaboration on how to accomplish securing devices and applications while allowing them to communicate as needed. C14 suggested tailoring or simplifying "cyber requirements" for the OT environment. E15 said, "You can't have them putting so many layers of protection that you can't use the system or operate it" and suggested "minimiz[ing] the complexity of cyber protections to avoid not being able to restore power due to cyber protections." They added:

> *To keep power flowing I have to do maintenance, operation on these devices, and replace equipment. [Cyber protections] can't be so difficult on the SCADA or communications sides that I'm unable to perform maintenance, repairs, replacements, and get the grid up. The system should not be so complex that we can't make it work with [them]. (E15)*

E16 said that a system or product might still "work very well" despite "poor code quality" and cyber SMEs' disapproval, suggesting that there could be more flexibility in restrictions for products that continue to work well.

C14 provided some insight into the technical problems of applying recommendations, such as installing Microsoft updates "in a plant kind of environment" or in "the electric sector":

> *They [energy OT SMEs] can't reboot their systems all the time. ... But if you want to change your authentication on one system, you have to update all the other systems that talk to it, in order to continue talking to it. ... The redundancy is more costly than in like a server farm. (C14)*

This suggests that they were familiar with how typical security measures like updates might disrupt operations.

#### 4.6.4 (d) Security by Design

Four participants (2 OT, 2 Cyber) recommended that OT systems be initially designed with computer security in mind, rather than focusing primarily on the operational functionality and addressing security flaws later on. C3 recommended bringing cyber SMEs into the development process at the software or board level, as security consultants, saying that some problems may be "a lot easier to fix in the beginning than to try and go back and patch it." E16 suggested that software developers "clean up the code" while considering risks and "abuse cases" when developing a product:

> *There's a whole concept of product secure development cycle, and if developers were just to understand and follow the principles within the secure product development lifecycle, then 99% of the issues that we have today would go away because the developers weren't just trying to make things work. They were trying to make them work securely, using good coding practices, looking at risks. (E16)*

E10 made the design recommendation for equipment companies to make "the right choices at the very beginning of their design process" such as putting in "hardware protection that protects their software."

C14 expressed skepticism about revising design processes, suggesting that both expert groups might stick to old habits or mental models, saying, "We would hope to do a better job of it, but … we're going to go back to what we know also when we redesign it." They also noted that some OT systems are "obsolete" and difficult to understand even for energy OT SMEs, and yet that "trying to set up the infrastructure from scratch" would be very costly.

#### 4.6.4 (e) Educating SMEs

Participants made suggestions for things to teach cyber SMEs (4 OT, 5 Cyber) and energy OT SMEs (1 OT, 3 Cyber). Topics suggested for cyber SMEs include: context, energy infrastructure, functional purpose, intended use, configuration, what it controls and is connected to, system requirements, energy OT need for access, maintenance and operation, that some devices are not practically exploitable in OT contexts, what is being targeted, what to prioritize (avoid rabbit holes), and impact on controls. Topics suggested for energy OT SMEs include: hardware and software vulnerabilities, risks and attack vectors, system capabilities, how to monitor what's happening and detect anomalous behavior, and what red teams or attackers are looking for.

Some participants specified that even within their expert group, certain skills were needed. C14 suggested that curiosity and willingness to learn were prerequisites for vulnerability researchers. C12 also said that even among cyber SMEs, it was rare to find people who were exceptional at finding meaningful exploits, saying, "Being able to find exploits is a skill, and not everyone has it. … And I would probably put myself in that category." C13 suggested that a "good cyber security person" should have experience exploiting vulnerabilities." E4 said energy OT SMEs should be familiar with things like technology misuse, vulnerabilities and historical exploits for OT equipment.

## 4.7 Discussion

Our research provides empirical insight into self-reported strategies, perceptions and suggestions of a group of interdisciplinary cyber SMEs and energy OT SMEs regarding vulnerability impact assessment in energy OT contexts. While we found that responses about information necessary to conduct an impact assessment (RQ1) were broadly similar across both groups of participants, some responses suggested major differences in the ways cyber SMEs and energy OT SMEs think about risks posed by vulnerabilities in energy OT systems. Differences appeared in their discussions of certain topics, like cyber SMEs' more adversarial and detailed considerations about access and exploits, and energy OT SMEs' holistic considerations about the overall system and disruptions in operations (Section 4.6.2). Differences also appeared in their perceptions of the two groups (RQ2), which align with prior work on differences between critical infrastructure security and traditional IT security approaches (see prior work in Section 4.4.1).

We first discuss the significance of the differences (RQ1 and RQ2) we found between the two groups (Section 4.7.1). We then discuss the interdisciplinarity of this group of participants and how their responses highlight the need for cross domain exchanges (Section 4.7.2). We also offer ideas for potential follow-on work related to cross-domain collaboration and for future work given the limitations of this study (Section 4.7.3). Finally, we make recommendations echoing the suggestions of participants (RQ3) to make systems more usable, and to develop more effective communication and collaboration across domains in critical infrastructure security (Section 4.7.4).

### 4.7.1 Harnessing Differences in Approaches

Participants' self-reported strategies, perceptions, and suggestions convey some relative differences in approaches to vulnerability impact assessment between energy OT SMEs and cyber SMEs. Finding differences within this interdisciplinary group is particularly insightful, as the differences highlight emphases and mindsets that can persist despite cross-domain experience. Considering that the goal of cross-domain interaction is not necessarily a complete skills transfer but rather to seek benefits from exposure to other methods and ways of thinking, the differences addressed in our work draw attention to approaches and perceptions that can potentially complement each other in building overlap in understanding risk for energy OT systems. Below we consider a few differences conveyed in participants' responses.

First, our study provides examples of cyber SMEs' considerations about gaining access to networks and resources, tracing paths across boundaries, modifying devices and their functionality, and exploitability. The overlap between these more adversarial considerations and energy OT systems may be useful for energy OT SMEs to understand.

Our study also revealed the more holistic emphases of energy OT SMEs on the overall system, potential disruptions in operations, and risk mitigation, which aligns with prior work suggesting that "OT practitioners" are primarily concerned with physical resilience and safety aspects such as "equipment damage and continuous supply of 'essential services' " [273]. Educating cyber SMEs about the overall system and their redundancies could help them avoid problems suggested by participants such as overestimating potential impact.

We encourage professionals and researchers to work to ensure that important aspects relating to risk and impact are transferred across domains.

### 4.7.2 An Interdisciplinary Group

Participants in our study had repeated exposure to the other discipline and worked with the other group at a research organization focusing on the energy sector. Given the interdisciplinary background of all participants and their similar broad-level responses, it appears that many participants had already built cross-domain awareness that allowed them to consider both computer security and energy OT issues when assessing impact.

In our interviews, some participants made clear references to experiences where they gained understanding from the other kind of SME. These experiences likely helped them develop a model for each group's skills, occupational motivations, and weaknesses, and we therefore hypothesize that many of their considerations sprang out of exposure to the other discipline.

While we expected the groups to diverge in their perceptions of each other, we were surprised to find that they had consistent views of both groups. Even when speaking of their own group, participants shared critical views of limitations or weaknesses. We did not see many instances of resentment or annoyance (the only times we noted this was when energy OT SMEs spoke about security policies that prevented them from working or slowed them down). Rather, negative perceptions usually indicated recognition of particular tendencies.

Yet, interdisciplinary experience did not appear to have equalized their knowledge base; they did not replace each other. Participants were aware of their own gaps in knowledge and where they might need to consult the other type of expert, and they recognized the strengths of the other experts. This was consistent with prior work conveying differences between operational and security professionals (Section 4.4.1).

Indeed, the two groups' specializations appear distinct enough to imply that experts will continue to need to come together to contrast their perspectives and build cross-domain understanding. The problem remains that interdisciplinary security in critical infrastructure contexts is not the norm; it is uncommon for energy OT SMEs and cyber SMEs to have access to each other. Resource constraints may also prevent companies from being able to build interdisciplinary teams or bring people together. Our study suggests that discussions or exercises across groups working in the same context will provide valuable insight. Researchers and industry professionals must seek ways to facilitate cross-domain exchanges.

### 4.7.3 Future Work

Below we describe potential future work building on this study, addressing its limitations, and expanding into other infrastructure contexts.

First, following our discussion above about interdisciplinarity, we encourage future work that develops ways to foster effective and scalable cross-domain knowledge transfer in energy OT contexts. For example, such work could consider vulnerability impact assessment approaches of energy OT SMEs lacking computer security experience and test the influence of interventions, such as exposure to training, educational materials or interdisciplinary interactions with a cyber SME, on participants' risk assessment considerations.

Additionally, given our small sample size, we encourage future work that investigates whether our hypothesis that the interdisciplinary nature of the group leads to similar general approaches to impact assessment holds at a larger scale. It is possible that our thematic strategy codes were not

able to sufficiently capture differences in impact assessment approaches for this limited number of participants. Future work could address these limitations by conducting a larger-scale study with these two kinds of experts to test whether there is a difference in approaches between the two expert groups, as well as between interdisciplinary and non-interdisciplinary experts.

Future work could also conduct interview studies with similarly sized interdisciplinary groups to see if differences are more pronounced when discussing different topics, such as what mitigations are acceptable, when to accept certain levels of risk, best patching practices, or who is responsible for given aspects of energy OT security.

Finally, we also recommend researchers explore building cross-domain understanding in different infrastructure contexts, such as healthcare, water, and transportation. Such contexts similarly require professionals to learn to operate highly specialized and complicated systems, such that adding computer security understanding to their job requirements can pose training and educational challenges.

### 4.7.4 Recommendations Building on Suggestions

There is a dire need for cross-domain collaboration in energy OT operational, training, and educational contexts, as it is not the norm for energy OT SMEs and cyber SMEs to work together, especially given the short supply of computer security workers. In their suggestions addressing collaboration between the two groups (RQ3), cyber SMEs and energy OT SMEs emphasized the continuing need for cross-domain communication and knowledge sharing among people who understand vulnerabilities and energy OT systems, as well as usable security and security by design. We echo participants' suggestions in our recommendations below.

First, we reiterate participants' suggestions to make energy OT systems more usable. As conveyed by participants, low usability security requirements can prevent engineers from effectively doing their work, or worse, encourage engineers and operators to use shortcuts that leave open vulnerabilities to be exploited. Usable solutions could include collaboratively developing security policies or designs that take into account operational needs such as continuous operations and the ability to restore power, simplifying or reducing human-in-the-loop computer security requirements that are too complex or burdensome for energy OT SMEs, and finding ways to update systems with minimal downtime.

Since one of the roadblocks to cross-domain exchanges may be organizational structure and assignment of roles, we also echo participants' suggestions to integrate teams. We recommend that companies and researchers investigate potential benefits of un-siloing workers and encouraging cross pollination of ideas. Walking through interdisciplinary contexts from multiple angles can help stakeholders develop holistic solutions that integrate diverse considerations.

As it may not always be realistic to integrate teams, given limited resources and labor supply, we also encourage the design and development of tools and interventions to help avoid wasting limited resources of potentially overextended operational engineering and cyber security staff. In addition to our call above for researchers to develop effective ways to acquire and apply cross-domain knowledge, we recommend that utilities and energy operators educate cyber SMEs and energy OT SMEs on the other group's objectives, how they think about a system or context, and information they might consider critical to understanding risk in operational contexts and computer security.

In particular, such training would provide energy OT SMEs with additional information on how to think about energy OT systems by considering different perspectives. Cross domain understanding could act as a companion to industry risk assessment standards, helping operators and other energy OT SMEs interpret standards with more nuance, rather than mechanically following checklists or output from automated systems, thus building resiliency in the human operators of energy OT systems.

## 4.8 Conclusion

We interviewed two groups of subject matter experts, energy OT SMEs and cyber SMEs, to explore and compare the two groups' self-reported impact assessment strategies, perceptions of differences between the groups, and suggestions for working together. We find that while their impact assessment considerations were generally similar, the details of their considerations and their discussions of their perceptions of each group revealed major differences in mindset and understanding. We recommend following participants' suggestions to foster interdisciplinary collaboration and integrate usable security into operational contexts, and we call for researchers and companies to develop tools and interventions that will enable cross-domain knowledge sharing in critical infrastructure security.

| Participant | SME Group | Experience | Prior Job |
|---|---|---|---|
| E1 | Energy OT | 11-15 | Y |
| E4 | Energy OT | 11-15 | Y |
| E6 | Energy OT | 21-25 | Y |
| E7 | Energy OT | 16-20 | Y |
| E8 | Energy OT | 11-15 | Y |
| E10 | Energy OT | 26-30 | Y |
| E15 | Energy OT | 31-35 | Y |
| E16 | Energy OT | 16-20 | N |
| E18 | Energy OT | 11-15 | Y |
| C2 | Cyber | 16-20 | N |
| C3 | Cyber | 1-5 | Y |
| C5 | Cyber | 1-5 | N |
| C9 | Cyber | 1-5 | N |
| C11 | Cyber | 16-20 | N |
| C12 | Cyber | 11-15 | Y |
| C13 | Cyber | 16-20 | N |
| C14 | Cyber | 21-25 | Y |
| C17 | Cyber | 1-5 | N |

Table 4.2: Summary of participants, showing participant number, expert group, total years of work experience (including prior experience), and whether or not they had work experience prior to working at the current organization.

| Differences | Description |
|---|---|
| *Cyber Focus:* | |
| Gaining access | Cyber SMEs discussed ability to move around within environments/networks and access additional resources to chain together |
| Identifying connections | Cyber SMEs wanted to trace paths across networks and systems, determine boundaries |
| Device capabilities | Cyber SMEs imagined potential capabilities such as misprogramming or modifying systems for different functionality, running arbitrary code, sending commands, or exploiting unused parts of hardware |
| Exploit details | Cyber SMEs considered exploit methods and also considered exploitability in terms of difficulty, not as a binary |
| *Energy OT Focus:* | |
| Connections to larger system | Energy OT SMEs were concerned with how the affected system connected to the larger system in terms of operations, important files, centralized SCADA/EMS systems, and downstream devices like smart meters |
| Disruption in operations | Energy OT SMEs specified considerations about disruptions in operations, such as whether the site was a high priority site, the amount of power at stake, and the severity of disruption |
| Risk mitigation | Energy OT SMEs emphasized containing the risk, ensuring operational integrity and investigating residual impact on the system |

Table 4.3: Summary of differences in vulnerability impact assessment strategies.

| Stereotype | Definition |
|---|---|
| Cyber does not understand energy OT systems | Cyber SMEs do not understand how energy OT systems work and may overestimate impact on overall system |
| Cyber sees computers | For Cyber SMEs, energy OT devices can be reduced to computers |
| Cyber is detail-oriented | Cyber SMEs pay attention to details, go into rabbit holes, spend a long time on analysis |
| Cyber cuts off access to protect system | Cyber SMEs place protections on the system that prevent or make it difficult for Energy OT SMEs to access or operate systems |
| Cyber equals IT | Cyber security and IT staff/departments are the same |
| Energy OT understands systems | Energy OT SMEs are skilled in the design, maintenance, and operation of energy systems |
| Energy OT does not understand vulnerabilities | Energy OT SMEs do not understand exploit capabilities, attacks, and details about vulnerabilities and may underestimate ease of exploit |
| Energy OT lacks imagination | Energy OT SMEs do not or cannot think of technical possibilities outside of what they already know, e.g., an adversary changing a device's functionality |
| Energy OT takes shortcuts | Energy OT SMEs leave access open, create backdoors, or otherwise allow vulnerabilities to remain, for convenience or increased usability |

Table 4.4: Specific stereotypes from responses comparing the two expert groups' strategies for vulnerability impact assessment and understanding of vulnerabilities. See Appendix C.4 for more thematic codes.

# Chapter 5

# Translation and AI Health Assistants in Healthcare Contexts

This chapter is an initial version of the following work in progress:

> Andrea Gallardo, Alexandra Li, Ray Liu, Elena Roldan, Veronica Lin, Lujo Bauer and Lorrie Cranor. 2025. "Interviews with Mandarin and Spanish Speakers on Existing and Speculative Medical Language Services"

## 5.1 Overview

This exploratory study explores the emerging technological application of user-facing machine translation tools and AI agents used in medical contexts by investigating the experiences, preferences, and concerns of stakeholders using a context-sensitive and interdisciplinary approach. Our work exemplifies our interdisciplinary approach through its design, execution, and data analysis. We contextualized our study by situating it in the use case of medical translation and interpretation and recruiting participants who reported speaking a language other than English as their primary language, having limited English proficiency, and seeking medical services in the same geographic area. We engaged participants in their primary language, conducting a multilingual study and multilingual data analysis, and below we present preliminary results for eight of the 31 participants in our qualitative study. Our findings reveal important tradeoffs between user preferences, trust attitudes, and the availability and performance of interpretation and translation technologies and services. Finally, we make recommendations for interdisciplinary future work at the intersection of human-computer interaction, machine learning, and linguistics.

## 5.2 Introduction

Language access is a prevalent problem in US medical institutions, many of which struggle to provide adequate translation and interpretation services to individuals who speak a language

other than English (LOE). Language services, such as certified medical interpreters, can bridge communication gaps between medical providers and patients. However, at certain stages of the healthcare experience, patients may need to rely on ad-hoc interpreters, such as family members, untrained staff, or translation apps on their personal devices. In such stages, such as scheduling appointments or checking test results, using personal contacts or commercial apps may compromise confidentiality of medical information.

Additionally, the rise of technology-mediated healthcare interactions has enabled patients to book appointments, access records, and communicate with providers via text-based patient interfaces on websites or mobile apps. Yet, patients with limited English proficiency (LEP) may struggle to take advantage of this increased access to information, as patient-facing interfaces are not always accessible in LOE.

The increase in the amount of electronic health information available to patients and medical providers has also led to an increase in data collection and the use of patient data. If translations of data consent forms and privacy notices are unavailable or are inadequately translated, LEP individuals may experience unequal opportunities to opt in or out of data collection or may not receive adequate notice regarding data storage and use.

While prior work has focused on challenges in multilingual healthcare communications with LOE/LEP patients [16, 49, 105, 237, 332, 410], there is limited understanding of LOE patients' trust and privacy attitudes and concerns towards different translation and interpretation modalities that mediate the transfer of medication information. We investigated participants' experiences with these intermediaries, and elicit their trust attitudes, preferences, and privacy concerns across different translation and interpretation options.

Our work is situated within current medical language access practices in Pittsburgh. We focus on the experiences and perspectives of Mandarin- and Spanish-speaking individuals with LEP who have previously required translation or interpretation during medical visits. We chose to interview Mandarin- and Spanish-speaking individuals to facilitate the recruitment of participants, as many US immigrants come from Chinese- or Spanish-speaking countries [414, 415], and to explore some of the issues facing LOE speakers of even high-resource languages, such as Mandarin and Spanish, when communicating across languages in medical settings.

Our project broadly has the following research questions:

- **RQ1**: What are LOE and LEP participants' default and preferred translation and interpretation options in medical contexts? When comparing options, what are some perceived tradeoffs between various modalities? How do their experiences inform their preferences?
- **RQ2**: What are LOE and LEP participants' perceptions, desires, and concerns regarding a speculative AI healthcare assistant that could translate text, interpret conversations, and provide explanations, personalized interactions, and assistance with tasks such as scheduling appointments?
- **RQ3**: How much do participants trust existing options? What are their privacy attitudes and concerns about these options?

Through our interviews with 26 participants (19 in Spanish and 7 in Mandarin) and our preliminary qualitative analysis of eight participants' preferences, experiences, and difficulties accessing medical translation and interpretation, we provide highlights of some results for each research question. We report notable considerations relayed by participants when expressing their prefer-

ences or comparing options and consider how participants' responses to hypothetical app features shed light on unmet needs that could or could not be addressed via technology, such as difficulties calling a medical office to schedule appointments. Finally, we highlight some responses related to trust and privacy concerns. We also provide recommendations for the design of technologies that can not only fill gaps in services but might also improve upon existing options to address unmet needs and persistent problems surfaced by our study.

Our findings reveal important tradeoffs between user preferences, trust attitudes, and the availability and performance of interpretation and translation technologies and services. For example, when discussing trust, some participants emphasized availability, accuracy, and reliability, as opposed to confidentiality or security, suggesting that for some participants, urgently and adequately meeting their basic need for access to information may take priority over privacy concerns. By presenting the experiences, preferences, and attitudes of LOE and LEP individuals regarding translation of medical information, this study contributes an understanding of how they use technology to meet their basic need to access critical information and where gaps or problems exist and may continue to exist if unaddressed by humans or technology.

## 5.3 Related Work

We provide background on US immigrants, individual rights to access medical information in non-English languages, and healthcare data privacy law. We also note some prior work on language access in medical settings for LOE and LEP patients, the use of machine translation in healthcare communication, connections between language barriers and poorer health outcomes, and related work on privacy taxonomies, medical disclosure concerns, and user studies regarding trust in technology and AI applications.

### 5.3.1 Demographic Background

Immigrants are a vital part of American society, historically, culturally, and economically, contributing hundreds of billions of dollars in state, local, and federal taxes annually [81, 107, 209, 281]. In 16 US states, immigration is crucial for offsetting population decline [50]. Immigrants also fill critical labor shortages, working in "manual labor-intensive occupations" that US citizens suggest they don't want [214], such as healthcare support, food preparation and serving, building and grounds cleaning and maintenance, farming, construction, and production, transportation, and material moving [114, 176, 281, 300]. Immigrants also increasingly bring a "brain gain," with more college-educated people immigrating in recent years and filling a considerable proportion of highly skilled roles in computer and mathematical, architecture and engineering, and life, physical, and social science occupations [33, 176, 281, 300].

Furthermore, immigrants play an important role in the healthcare workforce. One recent publication notes immigrants' contribution to the US healthcare system: "On the whole, foreign-born residents create a net benefit to the United States by paying more into the system than they receive in government-funded medical benefits. As workers, they fill crucial roles at every level of the health care system, from badly needed home health aides to surgeons and researchers" [54]. Additionally, another report notes that in the US "[a]bout one in six hospital workers are immi-

grants," including significant numbers of Asian (40%) and Hispanic (22%) immigrants, and they make up 27% of all physicians and surgeons in U.S. hospitals, 22% of nursing assistants, 16% of registered nurses, as well as 29% of building cleaning and maintenance workers and 20% of food prep and service workers at US hospitals [169]. Immigrants also provide critical services such as home health care for seniors and people with disabilities, making up over 40% of home health aides [299].

Most foreign-born people residing in the US are from Latin America or Asia. The top 10 places of birth, in order of estimated population, are Mexico, India, China (excluding Hong Kong and Taiwan), the Philippines, El Salvador, Cuba, Vietnam, the Dominican Republic, Guatemala, and Korea [414]. Since 2010, the top nine places of birth of the largest number of immigrants have been Mexico, India, China, Hong Kong, and Taiwan (the former three as one place), the Philippines, Cuba, Guatemala, the Dominican Republic, El Salvador, and Vietnam [415]. Thus, studying the availability of services in Mandarin and Spanish, as our study does, is relevant to language access policies in the US.

### 5.3.2 Policy Background

In the US context, institutions that receive federal funding, such as most large hospital systems, are required to provide meaningful language access under the Civil Rights Act of 1964, which prohibits discrimination based on national origin, including language. Our study is situated in the US context, where there are over 25 million individuals with limited English proficiency (LEP) [297].

Medical providers are also required to protect patient data under health data protection laws such as the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the US [53]. Additionally, with recent laws mandating access to personal data, including the US federal law requiring patient access to doctors' notes [72], patients may choose to translate these notes or seek to ask follow-up questions at home, where they may not be provided with translation mechanisms by the medical provider.

Thus, there is a need for user-facing language access options. In such situations, LOE and LEP individuals may choose to use popular off-the-shelf machine translation applications designed for individual use rather than for communicating healthcare information. These include Google Translate or DeepL, as well as LLM chatbots, such as ChatGPT. Yet, prior work has shown that current machine translation options may pose privacy risks and lack transparent disclosures about their limitations, introducing privacy compliance and liability problems [29, 103, 280]. As they were not designed to be used by medical providers to communicate patient health information, they are not required to be HIPAA compliant and thus do not provide the same data protection. Thus, histories of user translation queries may give away health information to people with access to the same account or device, and a lack of transparency of data collection and data use could result in a person's health information being stored on company servers, used in training data, or distributed to third parties.

While much prior work on improving language technologies in medical settings has focused on training language models for medical contexts, improving transcription and translations that medical providers offer, and protecting patient data, there is less focus on designing from a user-centered perspective to address privacy risks and transparency about limitations for widely avail-

able machine translation applications. Our work conducts essential need-finding on the disclosure preferences and concerns of LOE and LEP individuals across various medical translation and interpretation modalities.

### 5.3.3 Language Access in Medical Settings

Our work investigates the experiences, preferences, and attitudes of LOE and LEP individuals regarding language services, such as human interpretation and machine translation, in medical settings. We consider various language access modalities, including "gold" standard language services such as certified medical interpreters or professional translations, ad-hoc language providers like family or friends, and the use of machine translation and machine interpreting applications, such as Google Translate. Below, we describe prior work that sheds light on the provision of these language services and relevant problems, preferences, and perceptions of stakeholders.

In medical settings, the professional "gold" standard for interpretation consists of certified medical interpreters who can fluently interpret between patients and medical providers in person. In their literature survey of 28 articles about professional medical interpreters, Karliner et al. found that "use of professional interpreters is associated with improved clinical care more than is use of ad hoc interpreters, and professional interpreters appear to raise the quality of clinical care for LEP patients to approach or equal that for patients without language barriers" [190]. Interpreters can also serve as advocates, acting as cultural brokers and building trust and developing personal relationships with patients, as suggested by studies on the role of professional interpreters [17, 40, 63, 165, 167, 395].

However, research on medical interpretation shows that interpreters are underutilized [101, 164, 166, 195, 368], with interpreter use varying among physicians and nurses [166]. Some providers consider language services to be costly [179, 215, 260], despite research showing that the cost of language services "may be recouped through reduced testing, shorter visits, and better compliance with treatment and follow-up instructions" [260]. Additionally, quality may vary between in-person and video or phone interpreters, and remote interpreters may face obstacles to interpreting adequately, as they may not be able to pick up on gestures or other nonverbal cues [63, 242, 324, 428]. Finally, interpretation services may not be available to patients at certain stages of their patient visit, such as when they are searching for a provider or scheduling an initial appointment.

In addition to professional interpreters and translators, LEP patients sometimes depend on ad-hoc non-professionals to provide language access, such as bilingual caregivers or family members [286]. Prior work with LOE and LEP immigrant patients has considered the benefits and downsides of such non-professional assistance. Zendedal's interview study with 21 migrant patients found that "informal interpreters were expected to perform the roles of advocates and caregivers" [447].

Prior work has also considered problems LEP speakers face accessing translations of written information on medical patient portals [4, 19, 52, 58, 131, 222, 244, 293, 365], in documents such as discharge instructions [90, 248, 355] and in pharmacy communications [204, 307]. In their respective studies, Knecht et al. and Olenig et al. interviewed Spanish-speaking individuals regarding pharmacy interactions, and both studies found that language access was a problem, with some participants reporting difficulty accessing information in Spanish as well as poor cus-

tomer service and unfriendly responses to requests for Spanish language services [204, 307, 446]. Zargarzadeh et al. conducted phone interviews with pharmacists in charge at 552 randomly selected retail pharmacies throughout California (US) regarding multilingual labels on prescription medications and found that "approximately one-third of the sample reported either not having the capacity or choosing not to provide MLs routinely" [446].

### 5.3.4 Worse Outcomes and Less Access

It is critical for LEP patients, who have been shown to have worse clinical outcomes [51, 100, 198, 310, 316, 354], to access medical information and communications in a language they understand. Chu et al. found in a survey with 1000 participants (487 English-speaking, 256 Spanish-speaking, 257 Chinese-speaking) that "Spanish- and Chinese-language preference was associated with higher odds of negative health information-seeking experiences" [64].

However, prior work has shown that LEP patients are less likely to have equal levels of access to telemedicine. Rodriguez et al. compared the experiences of telehealth users with LEP and those with English proficiency (EP) and found that patients with LEP had less access to telemedicine and were more likely than those with EP to report worse experiences with video visits compared to in-person visits [354]. Pathak et al. found that non-English speakers had a decreased odds of portal activation and of having a video visit [316]. Our study explores LOE and LEP individuals' attitudes towards patient-facing machine translation tools as a potential means to bridge the language access gap.

### 5.3.5 Medical Machine Translation

As Machine Translation (MT) has increasingly been used in medical settings, its use by healthcare providers and patients has been a topic of study [98, 103, 196, 264, 312, 422], including the use of Google Translate in clinical settings [197, 280, 400, 416]. Zappatore et al.'s literature review of 58 journal and conference articles on medical MT "highlighted clearly how MT (particularly NMT [Neural MT]) is gaining strength as a helpful resource in the absence of professional translators/interpreters and how several studies pointed out its effectiveness" [445]. This reiterates earlier work's findings that "there is an urgent need for effective, accessible, and safe tools, such as translation technology, to facilitate everyday communication to improve health outcomes" [312]. However, they also found "[a]n unbalanced distribution between MT applied to clinical communication (n=21) and to health education (n=37)," suggesting that less research has focused on MT applications outside of these settings, including patient-facing tools.

Additionally, prior work has shown that general MT tools can be inaccurate and lack contextual awareness when applied to medical language [197, 280, 416]. For example, Khoong et al. evaluated the performance of Google Translate on discharge instructions and found that it lacked the accuracy and context sensitivity of certified medical interpreters [197]. They recommended, "Clinicians using [Google Translate] can reduce potential harm by having patients read translations while receiving verbal instructions; being vigilant about spelling and grammar; and avoiding complicated grammar, medical jargon (eg, fingerstick), and colloquial English." [197]. Thus, for certain procedures, medical professionals may need quality assurances that translations are accurate and meet human professional standards [125].

Vieira et al. suggest that "[i]nstitutional budgetary pressures as well as rudimentary or non-existent official guidelines all contribute to uninformed MT use" and recommend setting "robust standards regarding the situations in which MT use is and is not admissible" and promoting MT awareness and literacy in "contexts involving doctor-patient communication" [422].

While current MT and LLM accuracy and performance may have improved significantly in recent years, the consideration of context-specificity remains. In medical settings, the significance and meaning of phrases can vary depending on factors such as the speaker, context, medical history, and other relevant details. In their interview study with clinicians, Mehandru et al. emphasize the context-specific and conversational nature of medical patient-clinician consultations, challenging the notion that translation is a neat transfer of meaning, and suggest "reorient[ing] the goals of MT systems for clinical use from seemingly objectively optimizing translation quality between two texts, to designing for the overall quality of cross-lingual patient-clinician communication" [264].

### 5.3.6 Machine Interpreting

Machine interpreting, which combines MT and automatic speech recognition (ASR), faces similar problems as MT regarding accuracy and quality of translations, as well as problems with processing speech input. For example, prior work has shown uneven performance across speech varieties for some ASR systems, which may perform worse (produce a higher word error rate) in recognizing speech for non-standardized dialects and accents [56, 121, 208, 325, 366, 404] and languages that do not have a significant amount of existing data for models to be trained on ("low-resource languages") [340, 341]. Voice assistants have been shown to have higher WER for speakers of certain dialects or accents, resulting in failure to understand users' commands [34, 251].

Additionally, while machine translation post-editing is a standard quality control measure for professional translations that use machine translation, simultaneous machine interpretation permits no time for post-editing by translation professionals and domain experts, increasing the risk of harm from incorrect translations [445].

### 5.3.7 Emerging Mobile Health Technologies with Translation Capabilities

Work on emerging mobile health technologies suggest an increase in "patient engagement and personalization" features such as translations and "personalized care," as well as "AI-powered decision support" [381]. This includes technology that could help patients translate medical documents using approved translation and security methods. Solomou et al. developed a mobile application that "utilizes the EU eHealth Digital Service Infrastructure (eHDSI) OpenNCP for translating patient summaries and the FHIR Smart Health Links Protocol for secure sharing" and found that participants expressed enthusiasm and "frequently noted that such an application could greatly improve their health management and facilitate more effective communication with healthcare professionals" [380, 382].

Yet, LLP individuals may still face challenges adopting new technologies or understanding the meaning of medical information [320]. In their work on chatbots for "addressing the healthcare barriers of migrant workers" in Taiwan, Tseng et al. found that LLP participants suggested

that communication barriers to understanding medical jargon and "cultural barriers" based on biases regarding modern medicine may hinder them from taking advantage of emerging technologies such as multilingual chatbots. For example, despite being able to use Google Translate and searching for related health information using search engine, "translations of medical terms remain unfamiliar" [413].

### 5.3.8 Privacy

Relevant literature includes work on privacy taxonomies and self-disclosure to computer technologies, and disclosure to interpreters and healthcare professionals.

**Privacy Taxonomies.** Lee et al. adapt Solove's [383] and Shahriar et al.'s [371] privacy taxonomies based on "patterns of documented privacy risks resulting from AI's capabilities and data requirements" [225], resulting in the following privacy risks: Aggregation, Disclosure, Distortion, Exclusion, Exposure, Identification, Increased Accessibility, Insecurity, Intrusion, Phrenology, Secondary use, and Surveillance. Additional risks or harms from Shariar et al.'s taxonomy include: Inaccurate Decisions, Non-Compliance with Privacy Regulations, and Non-Transparent AI. Jakobi et al. developed a taxonomy of nine "user-perceived privacy risks"of using technologies such as "web browsing, voice assistants and connected mobility" [180]. In addition to the aforementioned Intrusion, risks from Jakobi's taxonomy include: Behavioral Manipulation, Discrimination, Health Impairment, Law Enforcement, Material Harm, Negative Effects on Social Status and/or Relationships, Societal Risks, and Socio-technical Environmental Variables. While these taxonomies have broad coverage, several privacy risks, such as Disclosure, Exposure, Inaccurate Decisions, Health Impairment (unfavorable health-related uses of data), and Non-Compliance with Privacy Regulations, are particularly salient in the context of our study, where translation technologies mediate highly sensitive and critical health information.

**Disclosure to Technology.** Prior work has suggested that self-disclosure preferences and content may vary based on whether disclosed via computer, paper or in-person interview. Early studies in this area found that more people preferred to disclose sensitive information using computer forms rather than via in-person interviews [431] and that open-ended responses in online forms were longer and less inhibited than those on paper forms [199]. In a study comparing 158 "unacquainted individuals" who met either in person or through computer-mediated communication, the latter group "exhibited a greater proportion of more direct and intimate uncertainty reduction behaviors" including more intimate questions and answers [411].

Researchers have investigated privacy and security concerns about chatbots [69, 95, 185, 229, 255, 261, 361, 442], User studies have found that health chatbots may facilitate the disclosure of sensitive information [192, 265], with some study participants suggesting they may be more comfortable talking about sensitive health topics with a chatbot [192], Yet, health chatbots also pose many potential risks to users. In their review of the "current landscape of LLM-based mental health chatbots," Yuan et al. identify "recurring ethical concerns such as privacy issues, potential biases, and user over-dependency" [444].

Some researchers have utilized language models to identify self-disclosure, hypothesizing that automatic detection of self-disclosed health data could provide benefits such as "earlier detection and treatment of medical issues" [417]. Given AI chatbots' potential to be an "effective platform to encourage self-disclosure of important personal information and honest description of experiences or concerns regarding a health issue," Chuan et al. created a chatbot "embedded on a colon health information website to conduct medical interviews with study participants" and ask them "embarrassing questions" and evaluated BERT and GPT-3 language models' ability to "understand the conversation content." They found that the LLMs were able to effectively classify the "level of detail and the presence of self-disclosure" in participants' responses, suggesting that such models could help chatbots "to respond in an appropriate manner" [66].

**Disclosure to Interpreters and Healthcare Professionals.** Prior work on patient disclosure to interpreters [40, 44, 191] and disclosure of stigmatized medical conditions [246, 315] suggests that individuals may have privacy concerns about disclosing sensitive medical or personal information to other people, such as interpreters and their family. However, there is a gap in understanding patient privacy attitudes and preferences regarding the disclosure of personal medical information to translation technologies, as well as how these compare to their attitudes about disclosing the same information to human interpreters or translators.

## 5.4 Methods

Below, we describe methods we used for recruitment, enrollment, interviews, translation, and data analysis. The study has been approved by our Institutional Review Board.

### 5.4.1 Recruitment and Enrollment of Participants

We recruited participants through our contacts in the local community, developing relationships with nonprofit organizations serving local Hispanic and Chinese communities. One nonprofit required an application to conduct research and a review of our materials, after which they agreed to share our call for participants. Additionally, we posted flyers at local food establishments and asked personal contacts to share our study information on WeChat, including a food vendor and WeChat groups based in the area. Finally, we asked participants to share the study information with their contacts.

We administered a screening survey to all participants to confirm that they were over 18, located in the US, went to the doctor at least once in the last year, spoke Mandarin or Spanish as their first language, and had limited proficiency in English. We asked participants about their proficiency in English using the Spanish and Chinese versions of the U.S. Census Bureau's American Community Survey questions that ask how well a person can speak and read English, where the answers options are: Very well (Muy bien, 非常好), Well (Bien, 很好), Not well (No bien, 不太好), and Not at all (No hablo inglés, 完全不会讲). We considered participants to have limited English proficiency if they did not select the "Very well" option for both questions. Thus, we had a few participants who qualified despite selecting "Well" for both questions. We also collected demographic information to help us diversify our participant pool in terms of factors such as age, income and education.

There was no compensation for participating in the screening survey; however, individuals who completed the screening survey were entered into a lottery to win a $50 gift card. After ending recruitment and closing the screening survey, we will conduct a drawing of every unique entry in the screening survey. Our survey screening questions are included in Section D.1.

Individuals who qualified for the interview study based on their screening survey responses were contacted to schedule an interview. We provided different ways to be contacted, including email, text message, WhatsApp, WeChat, and phone call. To communicate by email, we used an institutional Google account dedicated to the research project. To communicate by phone and WhatsApp, we used phone numbers dedicated to the research project. To schedule appointments, potential participants could provide availability through the above contact methods or use a Calendly online calendar booking page.

### 5.4.2 Interviews

We conducted semi-structured 75-minute interviews with individual participants. Interviews were conducted remotely and recorded and automatically transcribed via Zoom. Interviews were conducted in Mandarin and Spanish by members of the research team fluent in these languages. Participants were paid $60 for a 75-minute interview.

We designed our interview to elicit participants' strategies for accessing medical information in their first language and the challenges they encountered, and to document their opinions and attitudes regarding various interpretation and translation options as well as features of a hypothetical multilingual AI health assistant that could provide translation and interpretation and other potentially helpful features. Given the semi-structured nature of the interview, we did not ask every question in each interview, allowing us to explore various AI app features across interviews. The interview questions are included in Section D.2 and summarized below.

In the interviews, we gathered participants' prior experiences, strategies, trust levels, probability of use, and problems with existing interpretation and translation options, including human interpreters (e.g., remote or in-person medical interpreters, family members, or friends) and language technologies such as machine translation apps and AI chatbots. We first asked participants to describe the strategies they used to communicate with medical providers who only speak English, and the positive and negative aspects of the option they used most often. We also asked about prior experience with human interpreters and translation apps.

We then presented a hypothetical scenario in which they were asked to imagine a situation in which they just had surgery, received a doctor's medication instructions through spoken interpretation, then received written medication instructions to take home, and once at home, they realize that they didn't remember all the instructions and that all the documents are in English. We then asked them how they would go about trying to understand and follow the medical instructions. We asked them about how they would feel requesting documents from the medical office, using a translation app, and asking a friend or family member to help in this scenario.

We then described a hypothetical AI health assistant application that could provide translation and interpretation and elicited participants' initial impressions of the app. We asked them whether they would use this app for translating English medication instructions from home as well as for in-person interpretation at their doctor's office. We also asked how they would feel about potential features of this app, including explanations of medical information, appointment scheduling, informal conversation, region-specific accent and dialect selection, and personalized voice cloning.

Throughout the interview, we asked participants to compare options and express preferences among them for various medical contexts, such as interpreting at a medical visit, translating medication instructions at home, and contacting the medical office from home. Comparisons we elicited included comparisons between an in-person interpreter, a remote interpreter, a translation app and a hypothetical AI health assistant app. Towards the end of the interview, we also asked them to list uniquely human capabilities and human flaws related to interpretation and translation.

Finally, we asked participants questions about unmet language access needs in medical settings and suggestions they might have for a hospital that was developing the imaginary AI app.

### 5.4.3 Translations

We translated recruitment documents, the screening survey, and the interview script into Latin American Spanish and Simplified Mandarin using machine translation, via Phrase software, Google Translate, and DeepL API. Individuals outside of the research team who were fluent in both English and the target translation language conducted machine translation post-editing (MTPE) on

the translations to correct them. Final translations were reviewed by a member of the research team fluent in the target language. Following the interviews, team members fluent in Mandarin or Spanish reviewed and corrected the respective Zoom transcripts. Excerpts from transcripts were machine-translated into English using Google Translate, and qualified members of the research team conducted MTPE on the translations.

### 5.4.4 Data Analysis

Qualitative coding was conducted in multiple languages, with the first four authors, who also conducted the interviews, conducting qualitative coding of the interview transcripts. We first coded two interview transcripts in the original language of each interview, so that two authors coded in Mandarin and two in Spanish, with each pair double-coding two transcripts. After coding these transcripts, the four first authors discussed emerging themes in English, creating a multilingual code book in English, Mandarin, and Spanish. Team members fluent in Mandarin or Spanish provided language clarifications for non-fluent team members during discussions, when translations or meanings were ambiguous. For this thesis, we discuss preliminary results based on eight Spanish-language interview transcripts single-coded by the author, of which one was double-coded, as well as excerpts from other interviews.

To map responses to topics relevant to our research questions, we used a priori coding, labeling responses to certain interview questions or responses that mentioned certain topics using the following codes, with the respective research question in parentheses: Preference (RQ1), Comparison (RQ1), Human Traits (RQ1), AI Perception/Desire/Concern (RQ2), Unmet Needs (RQ2), Trust (RQ3), Privacy (RQ3). We also conducted emergent or bottom-up thematic coding, collecting themes that emerged from the data.

The naming convention we used for participants is based on the ISO 639-1 standard for two-letter abbreviations representing languages, where "es" represents Spanish and "zh" represents Chinese [235]. While "zh" is designated a Macrolanguage in this ISO standard, which "correspond[s] in a one-to-many manner" with many Chinese languages or dialects [234], we use "zh" since we only interviewed people in one Chinese language, Mandarin. Thus, participants are labeled with their corresponding language identifier and a number, e.g., ES11 or ZH03.

### 5.4.5 Limitations

This is a small-scale study and does not generalize. While a strength of our study is its focus on one city in the US, this limited scope means that we did not investigate patient experiences in cities or healthcare systems that may have significantly more or less resources, and thus our results are not broadly representative.

Additionally, our study is limited by our focus on high-resource languages and a sample of speakers who reported not knowing or who did not appear to know low-resource dialects or languages. This limitation prevents us from exploring how experiences, strategies and problems with limited access to interpretation and translation services may manifest differently for people whose primary language is not available or easily accessible via human language professionals or machine translation applications.

## 5.5 Results: Preferences and Comparisons (RQ1)

Participants' self-reported strategies for communicating with medical providers who only speak English included using in-person interpreters, remote interpreters, family members or spouses, translation apps, and AI chatbots. Below we present preliminary results regarding experiences and preferences that participants discussed regarding existing and hypothetical translation and interpretation options, including an AI health assistant app that could translate and interpret medical information.

### 5.5.1 Interpreters: In-Person and Remote

In discussing their prior experiences, participants who reported using human interpreters shared positive and negative aspects of working with them, which sometimes varied based on whether the interpreter was remote or in-person. Common challenges using either kind of human interpreter included inaccuracy, incompleteness, dialectal differences, lack of knowledge of medical terminology, as well as unavailability of interpreters. For example, ES09, who works as a nurse and who reported reading and speaking English "well" (see Section 5.4.1), emphasized interpreters' need for linguistic expertise and knowledge of medical terminology:

> *There are iPads that have the option to make video or phone calls. There are also physical interpreters, and almost always what I see is the same: the interpreters are from somewhere else. I've had some who are from Peru or Mexico, and the patient is from Puerto Rico or Cuba, and they use different words. Or they start using medical terminology, not the patient, but the doctors. And the interpreter gets lost. And the translation they try to give is wrong. So, obviously, there's no way to detect it if you don't know both languages and the medical terminology. But for me, who knows a little bit about everything, it's something I've noticed, and that's why I think that [knowledge of medical terminology] would be the most relevant thing for me in this situation. (ES09)* (Quote 1)

Multiple participants mentioned noticing that interpreters would leave out information they (participants) considered important. ES14 said, "sometimes you say something, and when the translator [interpreter] explains it, it's not what you said, or it's incomplete, so that you have to intervene and say, look, you failed to mention this information or you failed to say this."[1] When asked how they were able to recognize such failures to mention important information or inaccuracies, some participants responded that they had enough knowledge of English to recognize what was either being left out or interpreted incorrectly.

**Remote interpreters.** Challenges unique to remote interpreters included poor audio quality, loss of reception, and privacy issues, such as lack of consent for using a camera and exposure of personal information to other people in the remote interpreter's space, as discussed in Section 5.7.

ES09 recounted an experience with a supposedly disoriented patient who was actually having trouble understanding the remote interpreter via tablet because of dialectal differences and poor audio quality (low volume).

[1] See Appendix D.3 (Quote 2)

*I think he'd been in the hospital for two days, and that night I had to be with him. And so, from the moment I arrived, we spoke in Spanish and everything was fine. When they gave me the patient report, the nurse's report, they told me that the patient had been very disoriented, that he wasn't paying attention, and that he was very confused. When I spoke with the patient, the patient understood everything perfectly. And then one of the nurses showed me why. She brought the app [with] the interpreter into the room. We made a video call with someone from another country. And then he started asking the patient questions, and the patient answered incoherently, in the sense that, I don't know, they were asking him a yes or no question, and he answered something else. But in reality, the patient couldn't hear the iPad well. The volume was very low. And they also used terms that the patient didn't use. So, since he didn't know what they were asking him, he answered something else. And that's why everyone thought he was confused, that he was disoriented, that he didn't know what was going on. But in reality, it was a communication problem the whole time. (ES09)* (Quote 3)

Thus, ES09 witnessed an incorrect assessment of a patient based on poor quality and performance of interpretation technology and services. ES15 suggested that the wait for online interpreters could be long and the audio quality poor: "I have never really been able to call to make an appointment, because the vast majority always speak English, and when they offer online translation services, it is very very delayed. Sometimes you can't really hear the interpreter."[2]

**In-person Interpreters** ES07 recounted an experience with an in-person interpreter whose gaze and touch made a difference:

*Like when you're receiving difficult news or a difficult diagnosis. I experienced it, and on that occasion, the translator was there in person, and she felt and saw my pain when they were giving me the diagnosis. And it was very important to me that she simply looked at me with resilience or compassion. Although she didn't say anything to me, because she didn't have to tell me anything. She was an assistant. But she put her hand on my shoulder and told me: everything will be okay. And believe me, to this day, after 8 years, I haven't forgotten it. And I think if there had been a translator on, like, the blue phone [remote interpreter], well, he doesn't see my reaction, right? Or he doesn't see what's happening. And he couldn't have told me anything. The doctors, yes, in a certain way, in their protocols, but it doesn't even compare. I mean, the most they can say is "I'm sorry," right? But this person saw and felt my pain, and I felt good, and I feel like AI would never do that. No. (ES07)* (Quote 5)

Thus, when comparing this in-person interpreter to remote interpreters, doctors, and the hypothetical AI app, ES07 suggested that a human being dedicated to providing interpretation in person in a moment of crisis (a critical diagnosis) could most adequately read and respond to the pain she experienced in that moment.

**Translation Apps as Interpreters.** Participants who previously used translation apps to interpret during medical appointments suggested that it often made errors, sometimes stopped taking

[2]See Appendix D.3 (Quote 4)

in input, or might not work due to lack of network connection. To avoid transcription errors resulting in translation errors, ES02 said that when using a translation app to interpret her speech, she would verify whether the transcription was "exactly what I want to say" ("si es exactamente lo que quiero decir").

ES08 said she had not used a translation app as an interpreter in medical settings and preferred not to, because she found it impractical or unusable for longer conversations in which "time is running out":

> *When the conversations are so long, it's definitely easier to have an interpreter there. It's very time-consuming. You're there with your cell phone. Waiting for the person to tell you, read, delete, change the language because you have to change the arrow. And sometimes you don't change it because you're there, but time is running out, so you don't change it and you have to restart and repeat the message again and wait for them to read it, delete it again, change the message so they can talk to you. So it's very, very impractical in a moment like that, to do that whole process on your cell phone. (ES08)* (Quote 6)

Thus, usability, speed, and accuracy appear to be priorities and challenges for using translation apps as interpreters.

**AI app vs Existing Interpretation Options.** When we asked participants about using the AI app for interpreting at the doctor's office, participants weighed some of the tradeoffs, including potential risks of losing network connectivity, inability to recover from errors, and lack of human discernment. For example, ES09 suggested that AI apps would make errors because it could not detect sarcasm and might take things out of context: "Perhaps things can be taken out of context … especially if we make sarcastic comments or something like that. It's something that artificial intelligence can't detect. And so these are things that could be translated incorrectly. (ES09)"[3] Participants also highlighted potential benefits of immediate availability and quicker translations when using an AI app.

### 5.5.2 Translating Medical Documents

Responding to our hypothetical scenario about following medical instructions with only English language documents available, six participants said that they would use a translation app, with four of those specifying that they would use the function that allows them to take a photo of the documents and then see a translated version. ES14 said he would go back to the office to discuss it with the doctor, but that if this were not possible, he would use a translation app. One participant, ES09, said that she would first consult a family member and then use a translation app.

ES07 responded that in these situations, she took detailed notes, including in a situation in which she had to follow strict instructions to dispense medication to her child that would be dangerous in different doses:

> *I did this. I have this very present in my mind because one of my children has a condition, and his medication had to be exact. Every single detail had to be exact, so I wrote it down. And I still kept a note and everything, because I knew they were*

[3] See Appendix D.3 (Quote 7)

> *going to give me instructions for the medication. And when I got home, I mean, I couldn't take the risk because it was a very precise and very dangerous dose. Yes, so I had to translate line by line so I wouldn't make a mistake with his dosage and his schedule, because it was a daily medication that he has to take, and it was very delicate. Yes, yes, I've done it. (ES07)* (Quote 8)

Even though she took detailed notes at medical appointments, ES07 added that she supplemented these notes by using a translation app:

> *I've been through it several times, because ... when you're in the hospital, sometimes what they tell you, you're left with half of it, even if they tell you in Spanish. I'm very much a person who has to take notes. So, at that moment, I don't have a pencil and paper. So I say, Okay, but I'll make sure it's on the paper they're going to give me, the instructions. And as I said, [if I don't understand,] I translate it when I get home, and that's what I do, I translate it with an online translator. (ES07)* (Quote 9)

Thus, among our preliminary results, all participants suggested that they might use a translation app in this scenario.

### 5.5.3 Human Traits That AI May Lack

We share participants' perspectives on what human capabilities and flaws AI might lack.

**Human Traits: Positive.** Positive traits of humans provided by participants included empathy, as with ES07's emphasis on a compassionate gaze and touch, above, and personal contact or connection, i.e., "contacto personal" (ES14), and affinity or established bond (ES03). ES02 suggested that a human who saw her in really bad pain would recognize that she needs urgent medical attention, because a human could recognize anguish, desperation, or pain in someone's voice, while AI could not detect her crying ("llanto") or respond appropriately to the intensity that accompanied her expression of pain. She suggested that AI would interpret only the words that she said and thus not be able to map tone and emotion to a level of urgency.

**Human Traits: Negative.** Negative traits of humans provided by participants included being tired, sick, or moody, or allowing biased attitudes or beliefs to influence their work. For example, ES14 said that sometimes interpreters sound rude but also expressed compassion for their humanity:

> *Since we're human beings, we have moods, right? And we don't wake up the same way every day. Even if we want to, we still have problems. Sometimes I've noticed that some [interpreters] are kind of rude, and one has to understand that that person might be going through a bad time. I mean, they have to work because they have to earn a living, but we're human beings, so we have moods. So, moods often influence the service the interpreter provides. (ES14)* (Quote 10)

Thus, mood might affect human performance.

ES02 suggested that a belief that the patient was exaggerating could affect the translation, and that a human interpreter might translate "in the way they find convenient, not what I'm saying,

just say[ing] whatever he/she wants"[4]. She suggested that a translator might not take her seriously and thus produce a biased interpretation.

## 5.6 Results: AI Health Assistant App Perceptions, Desires, and Concerns (RQ2)

We present preliminary results about participants' perceptions, desires, and concerns regarding a speculative AI healthcare assistant that could translate text, interpret conversations, and provide additional features such as interactive conversations, explanation generation, agentic assistance scheduling appointments and requesting documents, dialect personalization, and voice cloning. As we presented results for participants' preferences and comparisons regarding an AI app as interpreter and translator above, we will focus on responses about three potential AI app features in this section: personalization based on dialect or accent, voice cloning, and agentic assistance contacting the medical office.

### 5.6.1 Personalization Based on Dialect or Accent

With this feature, a AI health assistant could understand and generate a dialect or accent chosen by the user. As shown in responses in Section 5.5, some participants discussed dialectal differences between the interpreters and patients interfering with the quality of translations in both in-person and remote interpretation contexts. For example, when discussing their experience with interpreters, ES14 said that "often, people who translate are not from the same region as us [...], so there are words we say differently," and that such lexical differences could "sometimes be inconvenient." Below, we present some participants' reactions to the potential ability of language technology to understand speech and generate responses in particular dialects and accents.

**AI App Understanding and Using Preferred Dialects and Accents.** ES06 suggested that developing a tool to handle various dialects could complicate things on a technical level in terms of accuracy and cause errors.

> *I think it would be great, but I'm also thinking, I mean, the programming part could get very complicated, and the fact that it gets complicated could cause errors. So, I think that, I mean, yes, as they say in Spanish, there are a lot of accents and everything, but when it comes to medical terms or medical translations, I feel like everything is somewhat neutral. You don't necessarily need a specific accent for Spanish speakers to understand. So, I think that with just normal, average Spanish, you could understand the entire application. (ES06)* (Quote 12)

Thus, ES06 suggests that the technical complexity of introducing dialectal variety could itself cause uneven performance across populations.

ES03 suggested that the joy of hearing the AI assistant speak in her accent would make her feel better:

[4]See Appendix D.3 (Quote 11)

> *Look, if it's the same as my accent, I'd be happy because it's like I'm in a doctor's office back home and they'd understand me: "Doctor: My head hurts, I have a yeyo." A yeyo, that's a sign of discomfort. "I'm dizzy." And if it answers me in my accent, that would be—I think I'd even feel better, from happiness. (ES03)* (Quote 13)

Thus, ES03 welcomed personalized interactions using regional accent or dialect and suggested it would improve not only the human-AI interaction, by being understood when she used a dialect specific term, but also improve her sense of well-being.

### 5.6.2 Voice Cloning

Reactions to the potential feature of setting the AI app's voice to be identical to one's own included "creepy" and "scary." However, a few participants said they would use this feature. For example, ES03, when asked how she would feel if the app had exactly the same voice as her, said "very happy" ("muy feliz"). When asked whether they would like the voice that the doctor is hearing to be the same as theirs or not, some participants' attitude reversed. ES14, who changed his view for this use case, said:

> *In that case, it would be different if I were to send a voice note, and the app would act like my voice, but for the doctor to hear it as my voice. That would be very different. That would be very good. Yes, I would like it. ... Because often you recognize people by their voice. ... So I think it helps communication. (ES14)* (Quote 14)

ES21, who said voice cloning would "scare" her ("me daría miedo"), suggested that this use of her voice could play a role in filling a gap in relaying emotion or sentiment in translation, reiterating her desire for AI assistants and human interpreters to imitate the "feeling" behind an utterance: "Yes, I think it would be more natural, but at first it would seem strange or odd to me, but I then think it doesn't lose what I was saying before [sentiment], because I think this application would basically eliminate that gap that I feel exists between interpreters. (ES21)"[5]

### 5.6.3 Appointment Scheduling and Contacting the Office

When we asked participants about having the AI app schedule appointments or request medical documents for them, participants suggested this would be a useful feature by discussing prior experiences, suggesting an unmet need. ES11 recounted the difficulty of calling a medical office: "I have been obligated to ask a family member to call for me, because I don't have another way to communicate with someone to ask for an appointment, if I can't understand them."[6] ES18 said that they had been hung up on for speaking Spanish when calling to make an appointment. ES14 expressed enthusiasm about using the AI app to schedule appointments, saying,

> *Just the act of making a medical appointment ... is complicated sometimes, due to the language issue. If I have an app that can help me ask for a medical appointment, that would be great. (ES14)* (Quote 17)

[5] See Appendix D.3 (Quote 15)
[6] See Appendix D.3 (Quote 16)

ES07 recounted an experience trying to obtain records from an office located in a basement and losing reception there, thus being prevented from communicating with the front desk until they found a security guard who helped them connect to the office's network.

Following our question about what she would do if given medication instructions in English, ES03 also rated the probability of asking the clinic for a translation as very probable (5), but when asked why she hadn't mentioned that option before, she said, "Because if I call the clinic, I wouldn't understand what they were saying, since they speak English."[7] This suggests that ES03 would use such a service if it were available and easy to use, but given current language barriers and usability issues, she would not realistically be able to make such a request.

## 5.7 Results: Trust and Privacy (RQ3)

We highlight some responses to our questions about trust and privacy. In our preliminary analysis, we found that some participants' privacy concerns varied depending on information attributes, such as whether it contained identifiable information or could cause them embarrassment. Additional disclosure concerns included criticality, suggesting health safety issues.

Though we anticipated that our questions on how much participants trust different interpretation or translation options would center on privacy, sensitivity of information, and disclosure, we were surprised to find that some participants' responses about trust were primarily concerned with the availability and quality of translation and the capability of the person or tool performing it, as discussed below.

### 5.7.1 Trust in Existing Options (RQ3)

Participants' reasoning for their trust rating answers sometimes focused on accuracy, completeness, and availability, rather than on trustworthiness of entities to safeguard their medical information. This was also reflected in the contrasting responses for probability of use with trust levels: sometimes participants rated a given interpretation or translation option with the highest probability of use and lower trust, or the highest probability of use with little to no prior experience using the option. For example, ES03 said it would be very probable that she would use an interpreter, because "everything I said, [the interpreter] said correctly," but rated her trust in an interpreter as a neither probable nor improbable (3), because "the doctor should not give information out to just anyone," saying that she might be concerned about her name and address being shared with the interpreter. This suggests the some participants may trust interpreters and apps to meet basic needs of communicating information but not, to the same extent, trust them to keep information private.

Indeed, some participants highlighted that their priority was understanding critical health information. ES04 emphasized the urgency of obtaining any kind of interpreter to handle urgent emergency situations, especially of her children:

> *If you go to an emergency room with a child. And if they don't have an interpreter, they have that computer to translate. As a mother, it really has to do with how they*

[7] See Appendix D.3 (Quote 18)

*treat your child, because you're not just going to leave it alone, because you don't know how to read English and don't ask for help. Your child is dying on you. What can you do? It's better to say, I don't speak English, can you get me a translator, please? Because they urgently need to see the child, What is happening? And then, if a person or a device or a phone is translating for you, what you want to see as a mother: that your child is okay. Because seeing your child sick, and you don't know what's going on. You want them to take care of him. (ES 04)* (Quote 19)

Here we see that a patient's hierarchy of needs may shift during a health crisis, where ES04's interest in understanding her child's health condition appeared to render concerns about data privacy secondary to her primary goal of immediately communicating with the provider. Thus, preferences, harms, and needs may operate dynamically, with the basic need for understanding critical information taking precedent in a hierarchy of needs.

Yet, this same participant noted that as her son gets older, she is increasingly left out at his appointments due to the language barrier:

*They speak Spanish there, but when you go to the pediatrician, he speaks in English. He says, if there's someone to translate, he'll translate for me. Otherwise, since my son is bigger now, he understands English. [The doctor] talks to him, he gives him his check-up and everything. I don't try anymore, I don't talk as much anymore, and [the doctor] uses the phone: "Everything is fine with your child and he's fine," he says, "There isn't a single concern. His weight is fine, his measurements are fine, his vision is fine." He uses the phone with me and talks more to my child, because they do this and that to him, they check his eyes, to read far away, everything [...]. I'm there with him. But now he just tells me, "Everything is fine with his growth." (ES04)* (Quote 20)

This suggests problems with translations apps could include quality problems as well as usability issues during appointments requiring interactions like physical exams, which might interfere with inputting text or speech on a phone or tablet. This lack of accuracy and completeness, could arguably result in a loss of trust, as ES04 may be lacking vital information or at least information that is very important to her.

In terms of trust, ES21 provided two ratings, "trust completely" (5) in terms of confidentiality, and "I am inclined to trust" (4) in terms of integrity, because she did not trust interpreters to accurately relay what she was saying, and even experienced an interpreter inserting their own speech into the translation:

*Sometimes there were things I would say to her, "I have this concern," but then I thought it would be better not to say it, because it would become a big deal. And on one occasion, I had an interpreter who also sort of scolded me. She told me, and precisely about that issue of baby bottle things, she would say, "Look, listen, she [medical provider] already told you not to give him that bottle."*
*And I would tell her, "It's not the bottle."*
*"You're not listening, she [medical provider] already told you not to give him another one."*
*So she didn't even interpret. She didn't even interpret or pass on the message. Rather, she would deduce what she [medical provider] had already told me, and then give*

*me the instruction. And then I would suddenly say to her, "But pass on the message that I'm telling you."*
*I started taking English classes and then I realized that she was no longer telling me what the person wanted to communicate to me, at the end of those sessions, which were, I don't know, maybe five. She said that I didn't drink milk, that I didn't consume dairy products. But I never said that. And then the person here, the one in therapy, always believed things that I wasn't doing and that I hadn't said. So it was frustrating. And that situation was chaotic. (ES21)* (Quote 21)

Thus, when ES21 discussed trust, she highlighted integrity and accuracy, having her message fully relayed, and the inappropriate insertion of the interpreter's own interpretation of the situation.

Privacy concerns about who can access information, data actors, data use, etc. did not appear to be front of mind when discussing trust for ES04, who similar to the excerpt above, emphasized that she had no other choice: "Trust completely, number 5, because really, I have two children, and if something serious is happening, and the interpreter is there, I have to tell her what's happening with the child. I can't not tell her, because how is she going to translate it for the doctor? (ES04)"[8]

### 5.7.2 Privacy Problems (RQ3)

Two participants encountered privacy problems in prior experiences with language services. First, ES01 said that while the remote interpreter was very friendly, the medical staff did not ask for her permission to keep the camera on: "They didn't turn off the camera, and they thought that was normal. [...] The medical staff turned around and left me alone with the interpreter. [...] The doctor was not looking at me. He was looking at the ceiling."[9] ES03 said she lost trust in interpreters after an experience with a remote interpreter, in which she could hear other people in the background near the interpreter, who was visible on camera via a tablet:

*There were people there. They could hear each other, and she was fighting [with them]. So there were more people listening to what I was saying. So my information, that is, what I was saying to the doctor, wasn't being protected. [...] She apologized to me when the doctor left. So it shouldn't have happened. From then on, I lost confidence/trust in translation [interpretation]. (ES03)* (Quote 24)

Thus, ES03's exposure of private information caused her to lose trust in interpretation.

### 5.7.3 Speculative Privacy Concerns (RQ3)

We previously discussed the potential opportunities and challenges for a multilingual AI health assistant to address existing needs and problems (Section 5.6). Here, we offer some highlights from participants' responses regarding privacy issues that might occur with this app, regarding access controls, data use, and resignation.

**Data Access: Desired Access Controls.** ES02 suggested using fingerprint or facial recognition authentication or locks to prevent someone with access to the same device from seeing someone's

[8] See Appendix D.3 (Quote 22)
[9] See Appendix D.3 (Quote 23)

entire medical history. ES03 emphasized secure storage to avoid access by unauthorized third parties, and suggested that app developers limit access for specialists to only what they need to see.

**Data Use: Voice Cloning.** As shown in Section 5.6, some participants found voice cloning to be acceptable, and a few changed their attitude based on whether the voice was intended for them or for the doctor to hear, suggesting that a personalized AI voice relaying what they are saying, including the sentiment behind it, might be desirable for some people. This suggests that the use case or intended purpose for cloning a patient's voice data may influence acceptability of this feature.

**Resignation.** When asked what his privacy concerns for the AI app were, ES14's response evoked overreaching data collection, but in the end, he stated that he was not worried:

> *Everyone tells you, yes, we keep your data safe here, ... but we're noticing that if you say something or you do a Google search, or you type something, and within two seconds, you're already seeing ads for what you were looking for. They're listening to us everywhere. Technology has, better said, accosted us. So, I think that the privacy of information, because the most valuable thing in this world is information, whoever has the information has the power. But I think that, honestly, many people have access to all our information. They're only missing access to our bank accounts, and I think they already have that too. So, that doesn't worry me anymore. Honestly. (ES14)* (Quote 25)

When asked subsequently if he would be concerned about who else might see his conversations with the AI health assistant, he replied, "To be honest, no. I think information is so widely available these days that, no, everyone has it now."[10] Thus, while ES14 provided a narrative of an omnipresent and powerful data collection status quo, in which data collection and use spans search, advertising, banking, and health contexts, and suggested that information is power, he nevertheless expressed a lack of concern and resignation that everyone ("todo el mundo") already has his information.

## 5.8 Discussion

Our study conveys how LOE and LEP individuals use or might use technology to meet their basic need to access critical information in medical contexts. Their discussion of current and potential issues reveals where gaps or problems exist with translation and interpretation, such as accuracy, privacy, and availability issues. Our findings reveal important tradeoffs between trust, preferences, and performance given such problems. We therefore provide recommendations for the responsible development of patient-facing language services technologies. We also provide future work recommendations for future research and work.

[10] See Appendix D.3 (Quote 26)

### 5.8.1 Designing AI Health Agents

We provide recommendations for the design of AI health assistants regarding trustworthiness and agentic assistance with administrative tasks.

Participants in our study contributed valuable insights for designing trustworthy user-facing applications for translation and interpretation in medical contexts. Their responses about trust and privacy (Section 5.7) suggest that such technologies should be available in emergencies and urgent situations (ES04), usable (ES04), accurate (ES21), and privacy-preserving (ES03). Additionally, reported negative prior experiences with dismissive or unsympathetic interpreters suggest that professionalism and empathy are important factors for building trust, even in non-human agents.

Going beyond translation and interpretation, our study explores additional features that a multilingual AI health assistant might offer, such as requesting documents or scheduling appointments. We found that participants expressed desire for such services or shared frustrating experiences encountering difficulties when trying to call the medical office to do these tasks (Section 5.6.3). Therefore, our results indicate that such a feature could help empower individuals to communicate in important medical information.

### 5.8.2 Supporting Linguistic Variation for AI Apps

Dialect-sensitive translation technologies could directly address the challenges reported by participants ES09 and ES14, who suggested that dialectal differences between interpreters and patients could impede understanding and lead to incomplete communication, including, as conveyed in Quote 3, misunderstandings about a patient's mental condition. We encourage technologists, language professionals, researchers, and other relevant stakeholders, including speakers of low-resource dialects and languages, to develop user-facing technologies that are dialect-sensitive and can adapt to different kinds of speech and accessibility needs, to address this persistent problem surfaced by our study, such that multi-dialectal AI interpreters might not only fill gaps in language access but also improve upon existing options. Despite these exciting opportunities to help patients better communicate with doctors, we also acknowledge existing performance disparities based on dialect for current LLM models, and discuss potential risks if such differences continue to exist.

**Challenges: Linguistic and Cultural Gaps.** Potential challenges to developing a multilingual and dialect-sensitive AI health assistant include problems accounting for lexical variants, semantic gaps, and relevant regional or cultural practices. Some participants' responses suggest that an AI system might not "know" some words, that is, that it would be unable to retrieve information such as dialectal variations of words or translations for medical terminology.

We anticipate that it will be challenging to develop robust language models as well as mechanisms for information retrieval to support AI agents that accurately parse and reproduce lexical variants or terms with the same meaning that vary across dialects, and that address semantic or lexical gaps between languages and dialects, with some terms having no easy or (sufficiently) one-to-one translation. Further work in NLP and computational linguistics that engages language professionals and fluent speakers is needed to document and validate language data in

NLP systems, e.g., including colloquialisms and dialectal variations, so that AI health assistants may adequately address the dialectal accuracy gap expressed by participants in our study.

Additionally, designing multi-dialectal and multilingual AI agents sensitive to linguistic and cultural variation could include developing AI systems that utilize a variety of cultural knowledge bases. For example, such systems could account for vocabulary related to relevant cultural practices, such as sports or traditional medicine common to a particular region.

**Risks of Dialect-Sensitive Technologies** ES06's skepticism about the technical complexity of implementing dialectal variation in AI agents and the potential consequent errors (Quote 12) raises important questions about how AI generated responses could be influenced by dialect. Not only might dialect adjustment cause differences in performance (WER, accuracy, information retrieval issues) based on how robust each version of the AI agent is for each dialect, but dialectal variation might also cause meaningfully different output, and therefore, materially different results.

Prior work shows worse LLM performance for non-standard dialects (see Section 3.4.3). Additionally, in recent prior work, Hofmann et al. show that LLMs can produce different decisions in high stakes contexts based on dialect, finding that "language models are more likely to suggest that speakers of AAE [African American English] be assigned less-prestigious jobs, be convicted of crimes and be sentenced to death" [156]. Furthermore, in recent work on "performance degradation," Lin et al. ran experiments with several LLM model families "covering four canonical reasoning categories: algorithm, math, logic, and integrated reasoning (tasks that require composing multiple reasoning skills)" and "anchoring these queries to known correct answers and employing human-based rewriting," and they found that "almost all models experience statistically significant performance drops on AAVE prompts, despite their semantic equivalence to their SE [Standard English] counterparts" [238].

Such dialect-related performance degradation suggests that technologists need to consider potential social and legal implications of dialect personalization, especially in healthcare contexts where algorithmic decision-making tools may reproduce suboptimal decisions disproportionately affecting populations who speak certain languages or dialects, or who have attributes that could be inferred from speech, such as disability or age. If patterns of suboptimal decisions are influenced by factors such as dehumanization, prejudice, or lack of resources, AI agents risk disproportionately providing low-quality services to certain populations and infringing on equal access rights.

### 5.8.3 Metrics

Technologists and researchers in ML and NLP are accustomed to working with metrics such as accuracy and efficiency, and in our study, we found that participants also prioritized factors such as accuracy, efficiency, and performance, as well as availability and privacy, when comparing translation and interpretation options. Yet, participants also noted other factors in their comparisons, such as emotional affect. Thus, our work exposes complex social considerations that demand alternate forms of accounting.

First, we recommend tracking thresholds for when users request or use translation or interpretation (which may vary individually), since we found that participants differed in whether,

when, and how they said they would use interpretation or translation services in medical translation contexts. Such metrics could inform the development of just-in-time language services, either human or machine. Furthermore, metrics on prior language services usage and availability across institutions, geographical regions and user demographics (e.g., health conditions) could also inform the provision of services.

Additionally, metrics on best practices for practicing sensitivity across various health situations and variables may be helpful for programming AI agents with a "bedside manner," or with simulated emotional intelligence. As we saw with ES21, who reported negative experiences with unsympathetic remote interpreters when facing difficult reproductive health issues, such as postpartum depression, emotional intelligence and sensitivity to particular health conditions is important to ensuring quality of service for interpretation. Similarly, metrics for wellness applications and technologies, such as features correlated with better or poorer health outcomes, could help inform the design of AI health assistants that provide personalized explanations of health information.

Finally, we suggest developing metrics for tracking linguistic features of AI agents, such as register or dialectal consistency. Such metrics would be distinct from accuracy, in that there may be various correct ways to say the same thing. Refining agents to consistently use and be able to switch between linguistic categories like formal or informal registers could improve the quality of AI agents.

### 5.8.4 Technology-Mediated Social Norms and User Choices

We expected more participants to recall interactions with in-person interpreters and for responses regarding positive aspects and interactions and informal conversation with interpreters to evoke themes found in prior work that suggest that interpreters act as advocates and helpful intermediaries. However, we were surprised to see that most participants had primarily interacted with remote interpreters, mentioned negative aspects of those interactions, and that some found informal conversation to be unnecessary (Section 5.5.1). Based on their negative experiences with technology-mediated communication, we see that technology is already impacting social relationships and norms between interpreters and patients. Thus, when comparing interpreters and AI agents, the tradeoff may be influenced by modality of human interpreters.

Additionally, participants' varying preferences and concerns about different interpretation and translation modalities, e.g., based on whether information is identifiable, embarrassing, or simple enough for them to translate themselves (Section 5.5 and Section 5.7), suggest the potential for tailoring language services based on patient preferences for a given modality based on sensitivity or complexity of information. These varying preferences suggest that privacy engineers should consider variable disclosure preferences, comfort levels, and designations of what constitutes sensitive information. Routing participants to a service with which they feel most comfortable relaying sensitive medical information will help ensure the most optimal communication between patients and doctors.

## 5.9 Conclusion

Our study on LOE and LEP Spanish-speaking and Mandarin-speaking individuals' preferences among translation and interpretation options shed light on persisting problems in language access as well as on important considerations for the provision of technology-mediated interpretation and translation, existing and hypothetical. Our presentation of hypothetical features to participants helped elucidate challenges, desires, and unmet needs that designers should consider as they develop patient-facing multilingual AI health assistants. We provide recommendations for patient-facing language services technologies and future work.

# Chapter 6

# Conclusion

Our context-sensitive approach considers factors such as population, location, use cases, sociocultural background, and language. Additionally, these studies are situated in particular times (of each study). Our interviews and survey elicited specific privacy, security, and performance considerations about the emerging technologies and contexts of: AR glasses, AI voice analysis, IT/OT convergence in energy systems, and AI health assistants. Our analysis uncovers important problems, existing and hypothetical, which must be considered in order to ensure that technologies are privacy-preserving, secure, and work well for end users with different needs and preferences. In our concluding chapter, we discuss some cross-cutting themes among the studies presented in this work, suggestions for future work, and policy implications.

## 6.1 Cross-Cutting Themes

We present two themes that apply across studies: contextual factors and user choice.

### 6.1.1 Contextual Factors

Our studies demonstrate the importance and value of contextualizing potential users or data subjects and use cases, and of scoping the problem or technology of interest, for investigating privacy and security problems posed by emerging applications of technologies. We consider contextual factors such as data subjects, data collectors, data recipients, information types, to echo parameters used in Nissenbaum's contextual integrity framework for aligning information flows with social norms. Furthermore, our work also considers sociocultural, linguistic and temporal contextual factors. While many different kinds of contextual factors, such as type of expertise, can inform the development of technology, it is impossible to scale such interview studies for every type of factor, and we therefore recommend prioritizing stakeholders who are highly relevant to specific problems.

**Expertise and Experience.** We engaged participants in our final three studies regarding their expertise and experience, i.e., their expertise in a particular domain, fluency in or familiarity with

certain US English dialects, and lived experience as native speakers of a language other than English (LOE) who have used language services in medical contexts. While our samples were small, we found that contextualizing their expertise and experience helped uncover important problems. In our study on subject matter experts' approaches to vulnerability impact assessment, participants' occupational motivations, influenced by their training and job experience, featured prominently and were important factors in their approaches. In our survey regarding AI voice analysis, some participants who self-identified as speakers of Appalachian shared perspectives informed by their linguistic background that reveal important implications for linguistic bias and discrimination. Finally, our study on LOE participants' experiences and perspectives on different interpretation and translation options conveys how critical contexts can reshape fundamental concepts like trust, where participants expressed trust in terms of whether interpretation or translation options could offer functional reliability and accuracy in critical situations, shifting meaning from data privacy to the ability to meet their basic need to understand information.

**Time.** In our work on medical translation and interpretation, participants' responses also suggested that time played a role in how they used language services, including increased understanding of English over time, which enabled some to ascertain the quality and accuracy of translation, as well as changing demand for language services based on what they preferred to communicate on their own. Furthermore, one participant also noted that as her son grew older and thus became able to communicate fluently in English with his pediatrician, the doctor consequently provided less complete translations to her and communicated more with her son, revealing how translation services might vary over time and age.

Additionally, attitudes towards technologies and privacy change over time, and our results documenting technology-specific preferences or opinions may thus have limited applicability. Yet, our results also illuminate priorities and concerns rooted in lived experience, social contexts, history, and linguistic factors that are not as malleable as technology features. These include fear of discrimination based on speech, disability, or language, desire for protections against unfair treatment by employers or schools, and suggestions for cross-domain understanding. Thus, while sociocultural challenges and opportunities remain, we hope that by giving voice to our participants, our work can helping address their needs by informing policy and technology design.

### 6.1.2 User Choice

Our studies on emerging speculative end-user technologies, AR glasses and AI health assistants, convey diverse and contradictory user preferences that suggest that these technologies should provide varying privacy, security, and personalization options that users can tailor. Personalization choices can help ensure the comfort of end-users, as some participants expressed discomfort regarding certain hypothetical features of AR glasses or AI health assistants, such as voice cloning.

Yet, while dominant paradigms of privacy suggest that user choice, notice and consent can serve as ways for individuals to be aware of and tailor the privacy and security features of technology, our findings regarding translation and interpretation services (Section 5.7) suggest that in critical situations, individuals may not really have a choice. In hiring and admissions scenarios, one can also imagine a coercive aspect to technology use by which candidates who do not

opt in to intrusive analysis would be excluded from consideration, as discussed in Section 3.6.4. Thus, designing for user choice and preference selection poses a challenge, as one would need to understand the most critical preferences while leaving space to refine less urgent preferences in non-critical situations. As conveyed in our work on medical translation and critical infrastructure security, when stakes are sufficiently high—whether concerning patient health emergencies or preventing a loss of power—the the necessity of urgent action, may leave end users with limited options to choose or to even understand the most private or secure options.

## 6.2 Policy Implications

We discuss policy implications and recommendations for policy makers regarding regulatory protections, evidentiary standards, and equal opportunity.

### 6.2.1 Protecting Against Unanticipated Consequences

The studies in this thesis consult potential end users or data subjects about the application of emerging technologies in various contexts. Indeed, many of the problems discussed by participants relate to potential unintended or negative consequences that technologies can have beyond intended or purported function or purpose and outside of the original contexts in which their underlying technologies were developed.

For example, in converging IT and OT environments, unexpected consequences can occur with the application of technologies or policies developed within one context without consideration for the other context. For example, in our interviews with subject matter experts (SMEs), participants' comments about rebooting devices reveal important differences that could result in unanticipated consequences if not considered. First, Cyber SME C11 suggested that energy OT SMEs overestimate the effect of reboots, because "they are used to things breaking and being good after a reboot or two," and that they may need convincing or explanations when told that a device has to be replaced due to a severe vulnerability (Section 4.6.3). Regarding updates and patches, Cyber SME C14 highlighted important unanticipated consequences of mandatory updates across an organization where "if you want to change your authentication on one system, you have to update all the other systems that talk to it, in order to continue talking to it." C14 suggested that when applying such policies "in a plant kind of environment" or in "the electric sector," the "redundancy is more costly than in like a server farm," because energy OT SMEs "can't reboot their systems all the time" and typical security measures like updates might disrupt operations (Section 4.6.4). Thus, our results show the need for contextualization and using knowledge from multiple domains to adequately asses potential risks to critical infrastructure.

Although some participants expressed enthusiasm or a lack of concern when considering speculative features of AR glasses, AI voice analysis, and AI assistants, this does not mean that there would not be risks or problems that are important for technologists and policy makers to consider. This includes norm-violating and overreaching uses of AI by employers who assess current or potential employees via AR glasses or AI voice analysis, without consideration for how those dialect and speech attributes can affect perception of ability or suitability for a job. Furthermore, AI translation software that does not account for necessary elements of translation

to ensure adequate medical communication could exacerbate or perpetuate existing problems with accuracy, availability, and trust.

Our work exploring AR glasses, AI voice analysis and AI health assistants suggests that such technologies may create social contexts and use data in ways that do not yet have established social norms. Such as-yet undefined norms demand reflection and consensus by stakeholders. Our studies showing a wide distribution of privacy concerns and considerations among current and potential users of emerging technologies also suggest that pre-emptive consideration for privacy preferences can help build trust. As we can anticipate technology companies testing the limits and legal boundaries of acceptable data collection and use, policy makers and researchers must actively preempt privacy harms, algorithmic harms, the violation of individual rights, and security and safety issues. We therefore call for policy makers to limit data collection, use, processing, and sharing by emerging technologies so as to align with reasonable expectations already pre-established in law or social norms for contexts in which such norms have existed before the emerging technologies. For example, as conveyed in Section 2.9, potential harms of features detecting or taking as input user attributes, such as disability or dialect, may include unintentional disclosure or false positives. We therefore also recommend privacy protections that prevent this data and its byproducts from being further distributed or used without a subject's consent, privacy notices such as privacy labels, and adjustable privacy options for end users and data subjects.

### 6.2.2 Evidentiary Standards

Given novel functionalities of emerging technologies, such as conversational AI-enabled information collection or biometric inference-making, we recommend formalizing information standards for AI-generated information and for information produced with the AI tools for supporting decision-making, especially in high-stakes contexts such as hiring, healthcare, and critical infrastructure. Indeed, such policy is already being proposed in the legal context, for evidentiary rules in U.S. courts, via Proposed Federal Rule of Evidence 707, which would demand that, "[t]o be admissible, the proponent of the evidence must show that the AI output is based on sufficient facts or data, produced through reliable principles and methods, and demonstrates a reliable application of the principles and methods to the facts" []. Such evidentiary standards are relevant in healthcare contexts as well, where an inaccurate AI transcription or translation might serve as part of a health record, which could have serious treatment, billing, and insurance consequences.

### 6.2.3 Equal Opportunity

In employment or education contexts, AI evaluations of current or prospective employees and students can have high-stakes consequences on income, future career prospects, grades, and admissions decisions. In considering difficulties encountered by individuals whose dialect or speech may be susceptible to algorithmic harms, participants suggested that AI voice analysis could violate their legal protections for equal opportunity, non-discrimination, and accessibility. Similarly, some participants imagined AR glasses social or conversational feedback negatively impacting their employment performance evaluations. As noted in Section 3.7, in the last few years, various countries and regional governments have passed regulations addressing this

problem [1, 2, 82, 120, 345, 348, 389, 434]. We therefore continue to recommend the further development of such policies to help safeguard people from AI-based discriminatory outcomes.

We also recommend that technology companies making AI evaluation tools for high-stakes contexts be legally required to retain each deployed iteration of its algorithmic evaluation tool, including training data and model settings, and that their clients (e.g., employers and schools) be required to retain all outputs, including AI-generated reports assessing data subjects. Such record-keeping will enable compliance with employment and equal opportunity laws, as well as robust discovery processes for litigation.

## 6.3 Future Work

We make suggestions for researchers and technologists, discussing potential future research and practices.

### 6.3.1 Interdisciplinary Research

Our work draws primarily from human computer interaction and usable privacy and security methods while including considerations from sociocultural anthropology, linguistics, sociolinguistics, and natural language processing. We recommend interdisciplinary human-centered work that further develops this approach, especially for for cultural and linguistics aspects of technology use.

For example, future work could further explore how technologists can prevent and mitigate potential harms of emerging technologies related to language and speech, caused by inaccuracies that affect performance or by harmful downstream effects of data processing, by developing end-user tools for flagging potential errors in real time, receiving transcripts of high-stakes interactions, and reporting issues to relevant quality control and compliance offices.

Additionally, as we suggested in Section 2.9, Section 3.7 and Section 5.8, we encourage technologists, language professionals, and researchers to engage with relevant potential end users and experts about developing user-facing technologies that are dialect-sensitive, can adapt to different kinds of speech and accessibility needs, and that also adapt to a variety of knowledge bases, such as traditional medicine.

### 6.3.2 Engaging Communities in Auditing

We also recommend auditing emerging technologies with diverse sets of stakeholders and call for auditing processes that engage stakeholders meaningfully. As Alondra Nelson has said, "It is not inevitable that AI will lead to great public benefits." AI that will benefit humanity should be "cultivated in partnership with civil society" and with "the meaningful involvement of the very people whose lives could be transformed by these technologies" [285]. We encourage technologists and researchers to engage with real people for auditing and avoid "diversity washing" data generation in which researchers attempt to account for under-consulted stakeholders by creating artificial data [435].

Philosophies such as contextual integrity emphasize the operation of technical systems within pre-established social norms. Our work helps document and surface expectations for systems, how they have operated, and how they should operate in social contexts. Auditing practices and collecting feedback from real people can thus help AI systems align with social norms by, as Nelson said elsewhere, "creating an expectation and benchmarks around what data you need … to actually even be able to pose the right questions about whether or not systems are operating the way they should," functionally and socially [454].

### 6.3.3 Whether Technology Is Needed At All

Our interviews and surveys with individuals elicited important problems and considerations that expose risks, challenges and opportunities of emerging technologies in high-stakes contexts. While the design and development of technologies should be informed by the needs, existing problems, desires, and concerns of stakeholders, meaningful consideration of stakeholders' interests must also include interrogating even the need for those technologies or whether a product should exist at all. Determining what kind of technology, if any, is needed or desired by stakeholders is the duty of researchers who are not bound by circumstance to develop products with predetermined functions, purposes and audiences. We hope that those with the ability to critically assess whether and how technology can serve people will publicly inform and develop innovative solutions (technical and non-technical) to existing problems.

# Appendix A

# Speculative Privacy Concerns about AR Glasses Data Collection (2023)

## A.1 Recruitment Text

Below are the texts we used to recruit potential participants: the initial recruitment text and the follow-up email to people who filled out the screening survey.

### A.1.1 Recruitment Text Posted to Reddit and Email Listserv

*Recruiting for Carnegie-Mellon University Study on AR Technologies: 90-minute remote interview, $30 compensation*

We are recruiting people with experience with augmented reality (AR) technologies for a research study. Each participant in the study will complete a remote interview over Zoom in which we will ask questions about your experiences with and thoughts about AR technologies. The study session will last approximately 90 minutes, and you will receive a $30 Amazon gift code as compensation after completing the interview. You may be eligible for this study if you meet the following criteria:

- You speak and understand English
- You are located in the U.S.
- You are at least 18 years old
- You can install and run Zoom for the interview
- You have used at least one augmented reality app or device recently. This could include any of the following, as well as other similar technologies: smartphone apps like Snapchat, Pokémon Go, Ingress, or Harry Potter: Wizards Unite or headsets or glasses such as Microsoft Hololens, Snapchat Spectacles, Google Glass, or Magic Leap 1

If you wish to participate, please complete our preliminary screening survey at [survey url]. If you are selected, one of our researchers will reach out to you for next steps. Thank you!

### A.1.2 Consent form email distribution text (appended to Qualtrics distribution message)

Thank you for filling out the screening survey for our study on augmented reality (AR) glasses. Our study will be a 60-minute interview, conducted over a Zoom video conference. We will ask questions about your experiences with augmented reality technology, such as smartphone games or augmented reality headsets. After the study, you will be compensated with a $20 Amazon gift code. If you are still interested in participating, please fill out the consent form and let us know when you are available by selecting an available time through the Calendly link at the end of the consent form survey.

## A.2 Interview Script

Below are the questions from our semi-structured interviews.

### A.2.1 Questions about current AR use

#### A.2.1 (a) Recent interaction details

- Have you ever worked in a job that required you to use AR?
- Have you taken a course that focused significantly on AR?
- In the last year, have you worked on any AR projects, for example for work, school, or as a hobby?
- What AR technologies have you used within the past year?
- Do you use an AR headset? Regularly?
    - What do you use your device for?
    - Where do you use it?
- Which AR app or device of those do you use most frequently?

#### A.2.1 (b) Current App/Device - Data Collection

- What data do you think [app/device] is collecting about you?
    - What do you think this data is being used for?
    - How do you feel about this data collection?

### A.2.2 General Attitudes and Expectations

- What would you like to do with AR glasses?
- What do you not want these AR glasses to be able to do?
- What things/data would you want AR glasses to collect/track?
- Is there any specific type of data you would not want AR glasses to collect or track?

### A.2.3 Data Types - Harms & Benefits of Data Use

We have a list of 15 types of data that your AR glasses could collect about you or your surroundings. I will first ask you how you would feel about AR glasses collecting each type of data. For some we may ask you to consider potential benefits or potential harms of AR glasses collecting and using this data about you.

1. Audio (Clarification: could be recordings, or could be always-on functionality, which would not store data long-term)
   *Benefit example:* Send voice commands to glasses, for example, to schedule an appointment while doing activities where it is inconvenient to use hands, like driving, cooking, or running.
   *Harm example:* Covertly record private conversations.
2. Video or Image data (Clarification: could be recordings, or could be always-on functionality, which would not store data long-term) *Benefit example*: AR glasses could scan your fridge and cupboard contents to make a list of items you need, to help you figure out what you need to get at the store. *Harm example*: Live streaming a social event without peoples' consent could reveal someone's personal information, like their family's images, to an audience they don't know and wouldn't share with.
3. Location data *Benefit example*: Navigational features like map directions. *Harm example*: Location data may reveal personal habits or relationships. Also, data mapped onto a virtual world could disclose personal information.
4. Maps of indoor spaces (Clarification: interior spaces of buildings) *Benefit example*: Furniture shoppers could see where a new piece of furniture would fit in their living room. *Harm example*: Prices and targeted advertising for online products could change because third party apps make assumptions based on home interiors, for example, based on objects you have.
5. Virtual spaces or virtuals locations that you have visited (spaces do not actually exist, e.g., VR Chat chatrooms, game worlds) *Benefit example*: A glasses user could keep a quick-access list of all the virtual spaces they like to visit so that they don't have to search every time. *Harm example*: Advertisements based on most-visited virtual spaces could pressure someone into making in-world add-on purchases that they later regret.
6. Your heart rate *Benefit example*: The glasses could alert users to heart conditions like heart arrhythmia, which could cause them to take steps to prevent potential health problems. *Harm example*: A record of heart rate activity could disclose patterns like sleep, exercise, and moments of anxiety or excitement. If someone shares this, it will reveal a lot of personal information.
7. Your body temperature *Benefit example*: Check for a fever to determine whether or not to leave home. *Harm example*: False positives, for example caused by using the device after being in contact with something hot, could result in automatic exclusion from events that screen for high temperatures.
8. Your brain waves (neural oscillations) *Benefit example*: Help detect any potentially harmful brain activities, such as strokes, seizures, sleep disorders, or other brain issues. *Harm*

*example*: Brain wave data aligned with visual data could be analyzed and potentially disclose personal info such as whether someone has seizures or make conclusions, such as recognizing someone.

9. Data about how you move, for example, gait (how you walk), posture, physical gestures, or body language *Benefit example*: The glasses could regularly remind users to realign their posture to help prevent long-term back issues and allow for gestural controls that work for users. *Harm example*: The glasses could read body language reactions to certain statements, people, or situations and make predictions about how the person will react to similar information in the future, making them more susceptible to targeted advertising based on their body language.
10. Eye tracking (your eyes) *Benefit example*: As someone reviews slides, art, or an event that they need to understand, they can see what they previously focused on (where their eyes looked) and discover things that they missed or barely acknowledged. *Harm example*: Companies might figure out what usually draws a person's attention and make conclusions about them.
11. Face images (remembering other people's faces) *Benefit example*: A glasses user can be reminded of someone's name as they walk towards them, just in time for them to offer a personal greeting. *Harm example*: Strangers will learn personal details about each other, which many people might consider invasive.
12. Facial expressions (yours and other people's) *Benefit example*: For someone using avatars or face filters, the glasses could sync their facial expressions with their avatar's face or a face filter. *Harm example*: Companies might analyze your facial expressions to see when you are most susceptible to marketing messages.
13. Your voiceprint, including tone and pitch of your voice *Benefit example*: The device owner's voice could be used to unlock device features, in other words, no one but that person could unlock the features. *Harm example*: The device or apps could share your voiceprint with third parties.
14. Other people's voiceprints, including tone and pitch *Benefit example*: Friends-only features could be unlocked in virtual spaces, where the space owner allows only certain people to access those features using their voices. *Harm example*: Deep fake applications that capture and imitate voices could enable people to fool others by sending audio messages pretending to be someone else.
15. Your reaction times, e.g. the amount of time you take to respond to a prompt *Benefit example*: Reaction times could be used to measure performance and help you improve performance, determine implicit attitudes, revealing unconscious biases, and action can be taken to avoid negative consequences. For example, a company realizes their hiring manager appears to automatically reject certain types of job applicants, so they change their practices to fairly evaluate such candidates. *Harm example*: Reaction times used to measure performance could result in burdensome consequences. For example, a student might be labeled a poor performer for slow reaction times, or a driver might be charged more for insurance for being slow on breaking. Alternatively could prove you were quick on breaks.

### A.2.4 Data Use

*For the following questions, provide a counterexample if their reaction is positive or negative and see how they respond.* Now we will talk about a few use cases that we came up with and ask for your opinion. There may be some overlap with what we just talked about, so please bear with us if there is some repetition.

1. How would you feel about notifications or reminders triggered by data being collected by the AR glasses, such as something you're looking at? *Benefit example*: a reminder that your pick-up order is ready as you walk by a store. *Harm example*: reminders could distract you when you are trying to focus or relax.
   (a) Do you feel this way about other reminders or notifications, like the ones on your phone or computer?
   (b) Would you like for your AR glasses to be able to predict what types of reminders would be helpful for you?
2. How would you feel about using AR glasses to monitor health, based on sensor data, like body and movement data? (*Benefit example*: Track vitals and detect dangerous irregularities before a potentially dangerous condition develops. Could get discounts based on positive data. *Harm example*: An app may share health information with insurance companies or the government, which may change rates or services based on data.)
   (a) Would you be willing to share AR glasses data about your health with doctors or other healthcare providers?
   (b) With researchers?
   (c) With a fitness tracker app?
3. How would you feel about using AR glasses for social or conversation feedback, such as how you speak, who you lean towards, or whether you interrupted someone? (*Benefit example*: Someone gets useful feedback on the tone of their voice as they practice a speech. *Harm example*: An employee gets negative evaluation from an AI tool that rates their conversation skills.)
4. How would you feel about setting a face filter in AR so that other glasses users could only see you with the face filter? (*Benefit example*: Make friends laugh with amusing filters. *Harm example*: Some people may get body dysmorphia and stop presenting their real faces, wishing to look like filters or avatars)
5. What do you think about AR glasses data being used to inform you about your mood or emotions? (*Benefit example*: During a conversation, someone gets feedback on their screen that they sound irritated, so they change their tone to sound calm, which leads to a more pleasant conversation. *Harm example*: Someone struggling to express themselves clearly becomes dependent on apps that inform them of their tone while they speak.)
6. Would you use the facial recognition feature? Would you allow other people with AR glasses to use facial recognition on you?
7. Based on the types of data we talked about, can you think of any other features you might expect AR glasses to have? Would you want this/these?

### A.2.5 General Questions - Data Collection

1. Does the location or where AR glasses collect data make a difference to you, e.g., whether you're out in public or at home?
2. Does the time of day, or when they could collect data make a difference to you?
3. Would data collection in certain social contexts make a difference to you, e.g. weddings, medical offices or places of worship?
4. Would you like to have control over who else the glasses could collect data about?
5. Would you like to know who or what companies would have access to data about you?
    - How would you like to be informed about that?
6. Would it make a difference to you if the data was stored locally, only on the device, on your personal cloud, on the company cloud, or somewhere else?
7. Does the length of time it is being stored make a difference to you?
    - *[If yes]* What length of time would make you uncomfortable for it to be stored?
8. Would you like to be able to delete your AR glasses data?
    - How important is this to you?
9. Would you like to be able to transfer your AR glasses data, e.g., from one brand of device to a different brand of device?
    - How do you imagine using this ability to transfer data?
10. When would you like to be informed about data collection? Before? While using? After? Some mix?
11. What would you like to know when you're being informed about data collection?

## A.3 Code Book

Below we include a copy of our code book containing the attitude codes and the emergent codes we used and developed throughout our qualitative coding process.

| Attitude | Definition | Statements/Keywords |
|---|---|---|
| Comfortable/support | Participant either said they would be comfortable, expressed enthusiasm with no hesitation, or expressed explicit support. If hesitation was expressed, ratio was 1/3 of enthusiasm or less, or expressed as an afterthought.<br>Qualified with Would Use. If coded Would Not Use, still expressed explicit support; otherwise label as Uncomfortable.<br>If they sympathize with negative, still express comfort or support personally.<br>Condition for non-use is very minor.<br>Compares to something they're already comfortable with.<br>Mentions it could be helpful, could be beneficial. | I'd be okay/comfortable with that. That would be cool. I don't really have any objection. |
| Conflicted/Mixed | Could see both sides, expresses equal amounts comfort and discomfort, could personally experience benefits and downsides without emphasizing either over the other. Context-dependent.<br>Contrasts two or more things, negative and positive balance. | I'd be uncomfortable with it, but would like it for X. I'm on the fence. It depends on... |
| Uncomfortable/oppose | Expresses discomfort, disagreement, negative feelings, concern. Points out potential problems. If they sympathize with positive, ratio was 1/3 or less. If they present a positive, seems like an afterthought (not emphasized).<br>Qualified with Existence not okay. If they say Existence Okay, they express explicit negative feelings; otherwise label as Comfortable<br>Condition for exceptional use case is very minor | I'd be uncomfortable with that. That's creepy. I wouldn't want it to. I don't see the point. I wouldn't use it, except in ... |
| **Attitude Emergent Code** | **Definition** | **Statement Examples** |
| Would Use | Mentions use case or says they would use it | I would use that for X. I'd love for {usage} to be a thing. |
| Would Not Use | Says they would not use it | I can't think of a use for it. I would not want that. |
| Existence Okay | Accepts other people using it or its existence despite their discomfort. NOT implied through "would use" or "comfortable" tag. | I would want it available to others. I could see how other people might use this. I'm not opposed to the technology itself. |
| Existence Not Okay | Objects to its use or existence | I don't think that data collection feature should exist. |
| Conditional | Provides condition for acceptance or use | So long as... Only if... |
| **Theme Emergent Code** | **Definition** | **Statement Examples** |
| Ability to Turn-off | Mentions ability to disable or turn off sensors or features | I want to be able to (manually) turn off this feature |
| Advertising | Mentions advertising or marketing | I wouldn't want to get ads while... |
| Amelioration | Improves quality of life for coping with disabilities. | It'd be helpful to me as someone with an ocular disorder/ADHD/autism/etc. This would be great for {condition} |
| Benefits Me | Says it would be useful, helpful, beneficial; provides utility | That would be useful for when I... |
| Bystanders | Considers situations, consent, or feelings of other people | I don't want it to record other people/my friends/my kids. |
| Collector Matters | Company or custodian of the data | It depends on who's getting my data. Do I trust the company? |
| Consent/Opt-in-out | Mentions consent or opting in/out | I want it to require my consent. I'd like the ability to opt-out. |
| Context/Situation Matters | Any situation (e.g., location, time) serving as a condition | Not in the home |
| Data Protection | Anonymization, encryption, PII revealed, secure storage | I wouldn't want my data to be tied back to me. |
| Data Use/Purpose | Concerns about how their data will be used: advertising, medical, product development, etc. | I'd be concerned about how they're using this data. |
| Data Retention Matters | Retention and deletion of data matters | I want to be able to delete or remove it |
| Some data off-limits/ Data Content Matters | What is collected (data content) or represented by the data matters. | It depends on what the data is (e.g., birthday, web history) |
| Discrimination | Unfair treatment, could exacerbate inequalities or social injustice | Employers could unfairly use this against me/others. |
| Initiated by User | Should occur only if user starts or requests it | Only if I start recording |
| I'm Used to It | Accustomed to data collection/usage as it already occurs on other devices. | I'm used to it. My phone already does this. |
| Legal Protection | Wants or imagines regulation compliance like HIPAA, GDPR, CCPA, etc. | I can't imagine that would be legal, given health laws. |
| Mental Health | Would have an effect on people's emotional and psychological well-being, ability to function in society, and meet the ordinary demands of everyday life. | I'd be worried about potentially worsening body dysmorphia or becoming over self-aware, addicted, or over-dependent. |
| Notice & Comprehension | Notification or some way to understand data policies is provided | I'd like to know what they're using my data for. |
| Personal Threat | Potential harm to person or property, e.g., theft, physical attack, identity theft | Data could be used by stalkers. |
| Recording | Whether device is recording, even for the short-term, affects how they feel. | I wouldn't want it to always be recording. |
| Storage Matters | Storage location of data. Local/device, personal cloud, company cloud | As long it's only stored locally on my device. |
| Third-party Access | Sharing data with third party companies, advertisers, employers | I don't want my employer to collect this data |

## A.4 Demographics

Our screening survey included questions about age, gender[1], race or ethnicity[2], and income. During our interviews, we gathered background information about the AR technology (e.g., device or app) they used most often and their attitudes[3] towards data collection for that technology.

[1]Gender: F=Female, M = Male, NB = Nonbinary, NL = Not Listed above, P = Prefer not to respond, A = Agender, G = Genderqueer

[2]Race/ethnicity: AA = African American/Black, AS = Asian, H = Hispanic/Latino/Latina/Latinx, W = White, NL = Not Listed

[3]Attitude: U: Uncomfortable/Opposed, M: Mixed/Conflicted, C: Comfortable/support

| Age | Gender | Race | Latinx | Income | Recruitment Source | AR Device/App | Attitude |
|---|---|---|---|---|---|---|---|
| 21 | P | NL | No | Prefer not to respond | r-augmentedreality | Pokemon Go | C |
| 31 | M | W | No | $50-60k | r-hololens | VR Device | C |
| 31 | NL | W | No | $20-30k | r-pokemongo | Pokemon Go | C |
| 30 | M | W | No | $100-150k | r-ingress | None | M |
| 58 | F | W | No | $100-150k | r-ingress | Ingress/HPWU | M |
| 24 | NB | AS | No | $80-90k | r-ingress | PolyCam | C |
| 25 | F | AS | No | $20-30k | r-pokemongo | Pokemon Go | M |
| 34 | NB | AS | No | $20-30k | r-ingress | Oculus | C |
| 33 | F | W | No | Prefer not to respond | r-ingress | Pokemon Go | M |
| 41 | F | W, AS | No | Prefer not to respond | r-hpwu | Ingress | U |
| 19 | F | B | No | $60-70k | r-hpwu | Instagram | No Response |
| 32 | M | W | No | $100-150k | r-hpwu | HPWU | C |
| 28 | F | A | No | $100-150k | nml | Snapchat / Instagram | M |
| 25 | M | NL | No | $80-90k | r-ingress | Pokemon Go | C |
| 31 | M | W | No | $50-60k | r-pokemongo | Hololens | M |
| 26 | A,F,G,NL | NL | No | $10-20k | r-ingress | Hololens | M |
| 27 | M | W | No | $100-150k | r-augmentedreality | Hololens | C |
| 27 | M | W | No | $100-150k | r-augmentedreality | Hololens | M |
| 24 | M | W | No | $50-60k | r-hololens | Hololens | M |
| 38 | M | W | No | $100-150k | r-hololens | Hololens | U |
| 46 | M | W | No | $100-150k | r-hololens | Google/Apple Maps | C |

# Appendix B

# U.S. Southerners' Attitudes Towards AI Analysis of Voice Data for High-Stakes Employment and Education Evaluations

## B.1 Surveys

We provide details about the pilot survey as well as the final text of the screening and main surveys.

### B.1.1 Pilot

In our pilot, we presented high-stakes use cases in a matrix question format, using the same Likert acceptability scale as in the final study (Section 3.5). Pilot results suggested that this kind of presentation was too brief and lacked context. For example, the education use case was described as "teacher correcting pronunciation," which resulted in a variety of interpretations. In order to reduce ambiguity in our questions, we made all use cases more specific. We also prefaced each voice-specific application with a general scenario to allow comparisons with algorithmic or AI applications not relating to voice data.

### B.1.2 Screening Survey

#### B.1.2 (a) Voice Assistant Knowledge

1. What do you think a voice assistant is?
   (a) A chatbot that generates responses to typed prompts
   (b) Computer technology that responses to spoken prompts
   (c) A service provided by hotels to make reservations or offer suggestions
   (d) A large language model
   (e) I don't know

#### B.1.2 (b) Voice Assistant Experience

1. Have you ever used a voice assistant?
2. Do smart phone voice assistants such as Siri on iPhones or Google Assistant on Android phones have difficulty understanding you?
3. Do home IoT voice assistants like Amazon Alexa or Google Home have difficulty understanding you?
4. Do you change the way you talk to be better understood by voice assistants?

#### B.1.2 (c) Region

1. What US state do you currently live in?
2. Were you born in the United States?
3. Did you spend most of your time before turning 18 in the U.S.?
4. (if Yes to the previous question) In what U.S. state(s) did you spend most of your time before turning 18?

#### B.1.2 (d) Language and Dialect

1. Is U.S. American English your first language? This can be any dialect or accent of American English, from any region in the U.S.
2. Is American English the language that you primarily speak?

The following questions are about the frequency with which you speak the following U.S. American English dialects or accents. *[Choice of: Never, Occasionally or sometimes, or Often or Always]*

1. How often do you speak General or Standardized U.S. American English (for example, the default American English dialect spoken by voice assistants, often used in American commercials)?
2. How often do you speak Southern U.S. American English (also called a "southern accent")?
3. How often do you speak African American English (also called African American Vernacular English or Black English)?
4. How often do you speak Appalachian English (commonly spoken in central and southern Appalachia, e.g., in West Virginia and Kentucky)?
5. Do you speak another dialect of U.S. American English not listed above? If yes, please provide the name of the dialect. If you speak multiple additional dialects, please provide the one you speak most often.
6. (if Yes to the previous question) How often do you speak the dialect you specified above?

#### B.1.2 (e) Speech

The following questions are about your speech, or how you utter words aloud.

1. Do other English speakers have difficulty understanding you when you speak? *[ Choice of: Never, Occasionally, Sometimes, About half the time, Most of the time, Always ]*
2. Do you have a speech disorder or voice disorder that could influence how well voice assistants understand you?
3. Are there other factors besides accent, dialect, speech disorder, or voice disorder that could influence how well voice assistants understand you? If so, please specify.

#### B.1.2 (f) Demographics

1. What is your approximate age in years?
2. How do you describe your gender identity?
3. How do you describe your race or ethnic identity?
4. Do you identify as Hispanic or Latino/Latina/Latinx?
5. What is the highest level of school you have completed or the highest degree you have received?
6. What was your approximate household income in 2023?
7. Do you have a formal education in a computer-related field, such as computer science or IT?
8. Do you have a formal education in a data science related field, such as statistics or machine learning?

### B.1.3 Main Survey

#### B.1.3 (a) Data Collectors

**Scenario:** Imagine that the following people or entities are collecting voice data from voice assistants that you encounter, for example, in a store or office. The entities will use your voice data to improve their services, product, or performance.

**Question:** How acceptable is it for you to give your voice data to the following data collectors? *[Choice of: Extremely unacceptable, Somewhat unacceptable, Neither acceptable nor unacceptable, Somewhat acceptable, Extremely acceptable]*

1. Voice assistant companies like Google, Apple, and Amazon
2. AI technology companies like Open AI, Google, Meta
3. My school district
4. Potential employer
5. My current employer
6. My teacher

### B.1.3 (b) Inferences

**Scenario:** Imagine that voice assistants can be programmed to infer or guess the following speaker attributes. That is, they could use a speaker's voice data to estimate attributes like age, gender, etc.

**Question:** How acceptable is it to you for voice assistants to make inferences about the following characteristics based on your voice data? *[Choice of: Extremely unacceptable, Somewhat unacceptable, Neither acceptable nor unacceptable, Somewhat acceptable, Extremely acceptable]*

1. Age
2. Dialect or accent
3. Ethnicity or race
4. Gender
5. Health condition
6. Mood or emotional state
7. Region of origin (where you grew up)
8. Sexual orientation
9. Speech or voice disorder

Next, we will ask you about four (4) scenarios related to technologies that could be used in education and employment settings, as well as potential applications of voice technologies in those settings. For each scenario, we will ask you to provide free-text responses. We're especially interested in your attitudes, opinions, or experiences relating to the southern United States and dialects or accents found in this region. Thank you for your thoughtful responses.

### B.1.3 (c) Education Admissions

**General Scenario** A small liberal arts college uses software to determine eligibility for admission to their undergraduate program. It takes application materials and ranks students based on admissions criteria.

1. How acceptable do you find the scenario above?
   (a) Extremely unacceptable
   (b) Somewhat unacceptable
   (c) Neither acceptable nor unacceptable
   (d) Somewhat acceptable
   (e) Extremely acceptable

Next, please consider the following voice-specific application of the technology from the scenario above.

**Voice Tech Scenario** The software evaluates audio recordings of admissions interviews with applicants to determine their suitability for small classes with interactive discussions. Students

will be evaluated based on ability to respond to questions using appropriate language and clearly articulate their opinions.

1. Does the inclusion of this voice-specific application make the scenario more or less acceptable, or make no difference in acceptability?
   (a) Less acceptable
   (b) Makes no difference in acceptability
   (c) More acceptable

**Free text responses**

1. You selected *(general scenario answer choice)* for the general scenario and *(voice-specific answer choice)* for the voice-specific application. Please explain your overall attitude for this scenario and voice-specific application.
2. If you believe that there are potential harms for this use case of voice technologies, please list at least one potential harm that you can imagine related to the above use case. If you do not believe that there are potential harms, please explain why.
3. If you believe that there are potential benefits for this use case of voice technologies, please list at least one potential benefit that you can imagine related to the above use case. If you do not believe that there are potential benefits, please explain why.

### B.1.3 (d) Education Evaluation

**General Scenario** A high school district uses automatic grading software to grade students' English annual final exam, which counts for 20% of their final grade in the course for the year (grade shown on transcript).

**Voice Tech Scenario** Exams include a spoken component, and the software grades students' English word pronunciation. The spoken component counts for 25% of the final exam grade.

### B.1.3 (e) Employment Hiring

**General Scenario** A company hiring new employees uses software to rank the top job applicants throughout the job application process. Employees are ranked based on past experience and skills. The hiring manager will review only the top 10 candidates on the list created by the software.

**Voice Tech Scenario** The software is used during in-person job interviews. It analyzes the job applicant's speech during an exercise in which the applicant pretends to speak to a customer and rates how understandable the applicant's speech will be to customers.

### B.1.3 (f) Employment Evaluation

**General Scenario** A company uses software to evaluate current employees' job performance to determine pay raises. Employees are told ahead of time about the metrics to be considered by

the software and can request re-evaluation for certain components that they can re-do, such as submitting their work product for review.

**Voice Tech Scenario** The software considers recent presentations by employees. It analyzes speech during presentations to evaluate public speaking and communication skills.

### B.1.3 (g) Experience and Usage

Voice Assistant Experience

Thank you for your written responses. In this final part of the study, we will ask you some questions about your use of voice assistants.

1. Which of the following devices do you use? Select all that apply and please specify the device name if you use a home voice assistant or other smart device (for example: Amazon Alexa or Google Home). *[ Choice of: Home voice assistant, Smart phone, Smart TV, Smart watch, Wearable fitness tracker, Other smart device, None of the above ]*
2. Do you use a voice assistant on your smart phone, such as Siri or Google Assistant? If yes, please specify which one(s).
3. Some voice assistants are phone-based. You might encounter them when calling a bank or an airline by phone. Have you ever encountered a phone-based voice assistant?
4. How often do you use voice assistants or other speech recognition technology for tasks such as dictation, searching, or controlling devices? *[Choice of: Rarely or never, Once every few months, Once every few weeks, Once per week, Daily ]*
5. How often are you around someone else's voice assistant, for example, a family member or roommate's voice assistant? *[Same choices as previous question.]*
6. Some voice assistants offer the option to improve their ability to recognize you by doing some short exercises to teach them your voice. Have you ever trained a voice assistant on your voice in this way?
7. Where do you use voice assistants? Please select all that apply. *[ Choice of: My home, Work, Car, Gaming, Home of friend(s), School, Other, NA ]*

Please indicate your level of agreement or disagreement with the following statements. *[Choice of: Strongly disagree, Somewhat disagree, Neither agree nor disagree, Somewhat agree, Strongly agree, N/A - I do not use voice assistants, or I don't know]*

1. Voice assistants usually understand me.
2. I engage less with voice assistants due to past interactions with voice assistants.
3. I am able to use voice assistants equally as well as the average or standard user they're developed for.
4. It is important to me that speech recognition technologies, such as voice assistants, accurately understand my individual voice.
5. It is important to me that speech recognition technologies, such as voice assistants, accurately understand voices that sound similar to mine, for example, with the same accent and dialect.

6. (Optional) Please add anything else you would like us to know about your experiences with voice assistants.

## B.2 Appendix - Code Book

Below, we include our code book with the thematic codes we used to label participants' responses regarding their overall attitude to each scenario for four use cases in Table B.1.

| Code | Subcode | Definition |
|---|---|---|
| Negative opinion (Critique) | Scenario critique | Critique of scenario details rather than voice tech application or decision-making AI, problem with the outcome variable or the fact that such a scenario would exist |
| | Discrimination or bias | Unfair or negative outcome due to speaker attribute(s); mentions discrimination, bias, unfairness, prejudice, etc. |
| | Unfair outcome | Data subjects are treated unjustly or receive an unfair outcome (rejected, filtered out, or evaluated unfairly) |
| | Privacy or security | Harms related to data privacy concerns, malicious use of data, security breaches, user consent to data use, etc. |
| | Errors or glitches | Errors or glitches of the technology will be harmful |
| Positive opinion (Support) | Adequate measure | Technology will be reliable, able to get the job done, capable of performing the given task adequately or better than humans |
| | Efficiency | AI or software would be efficient and save time, labor or resources, e.g., reduces workload or streamlines process. |
| | Mitigate human bias | AI or software could help mitigate potential bias of humans involved in decision-making process. Positive view of technology related to objectivity and consistency of technology. |
| | Improve skills | Evaluation will help data subjects improve their skills |
| | Important to stakeholder | Highlights the importance or relevance of the evaluation to stakeholders (company, employee, customer, school, student) as justification for the acceptability of software/data use |
| | Evaluation or Finds the best | Technology will help find the best or a better selection of job candidates or students, or to evaluate communication skills or pronunciation to select the best performers |
| Question or concern | Already happening | Notes existence of or giving examples of similar scenarios that are already happening |
| | Depends on data type | Software using some data types is OK, but not others |
| | Question or concern | Question or concern regarding scenario that they would like more information about or a condition or variable that would influence their response |
| Humans & Tech | Preference for human evaluation | A human should be involved in the process, do the task or review the job done by the software. People deserve review by humans. |
| | Inadequacies | AI, software or technology in general is not able to do the given task, either at all or as well as a human could |
| Speech-related comment | Dialect, accent, or speech | Response mentions dialect or accent, or comments on speech attributes like "how people speak" or differences in speech patterns and pronunciation |
| | Speech is not a proxy for intellect or skills | Speech or voice data does not adequately represent a person's abilities or intelligence and should not be used to gauge skills |

Table B.1: Thematic codes developed for responses to questions about participants' self-reported overall attitudes and potential benefits and harms towards software to assist decision making in educational and employment contexts.

## B.3 Appendix - Demographics

Participants self-reported their dialect, age, gender, race or ethnicity, and income, as shown by percentage in Table B.2. Note that multiple gender and racial or ethnic identities could be selected, making their respective column total more than 100%. When reporting speaking dialects often or always, participants reported various combinations of dialects, as shown in Figure B.1.

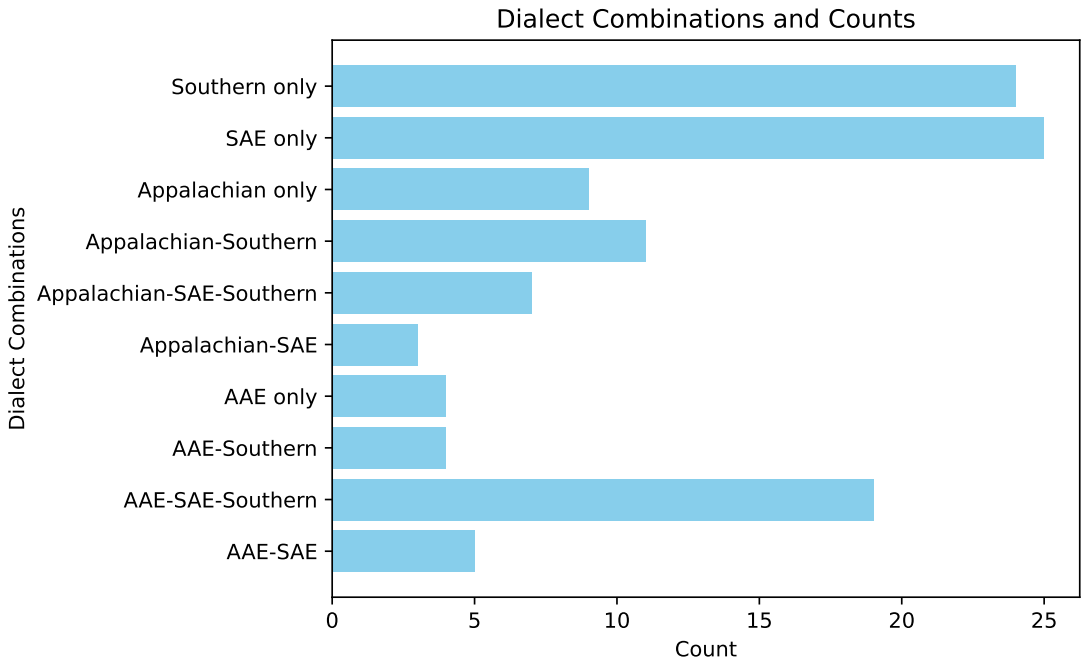


Figure B.1: Counts of dialects spoken by participants.

Through purposive sampling, we recruited participants in six groups: AAE, Appalachian, SAE-AAE, SAE-Appalachian, SAE-only, and Southern-only, whose definitions are provided in Table B.3. The AAE and Appalachian groups included participants who identified as speaking AAE or Appalachian often or always, respectively, including those who identified as speaking Southern often or always. The SAE-AAE and SAE-Appalachian groups included participants who identified as speaking AAE or Appalachian often or always, respectively, as well as SAE often or always (some also identified as speaking Southern often or always). No AAE or Appalachian speakers reported also speaking the other dialect often or always. The SAE and Southern groups each include participants who only identified as speaking SAE or Southern often or always, respectively, and exclude participants who identified as speaking any of the other three dialects often or always. As we were not able to recruit sufficient numbers for the AAE and SAE-Appalachian groups, we consolidated the six groups into four groups corresponding to the four dialects of AAE, Appalachian, SAE, and Southern, which we describe in Table 3.1.

We also gathered background information about participants' experience with and their attitudes towards voice technologies they used most often. When asked whether other English speakers, smart phones, or home IoT devices have difficulty understanding them when they speak, most participants responded "Never," "Occasionally," or "Sometimes," as shown in Figure B.2.

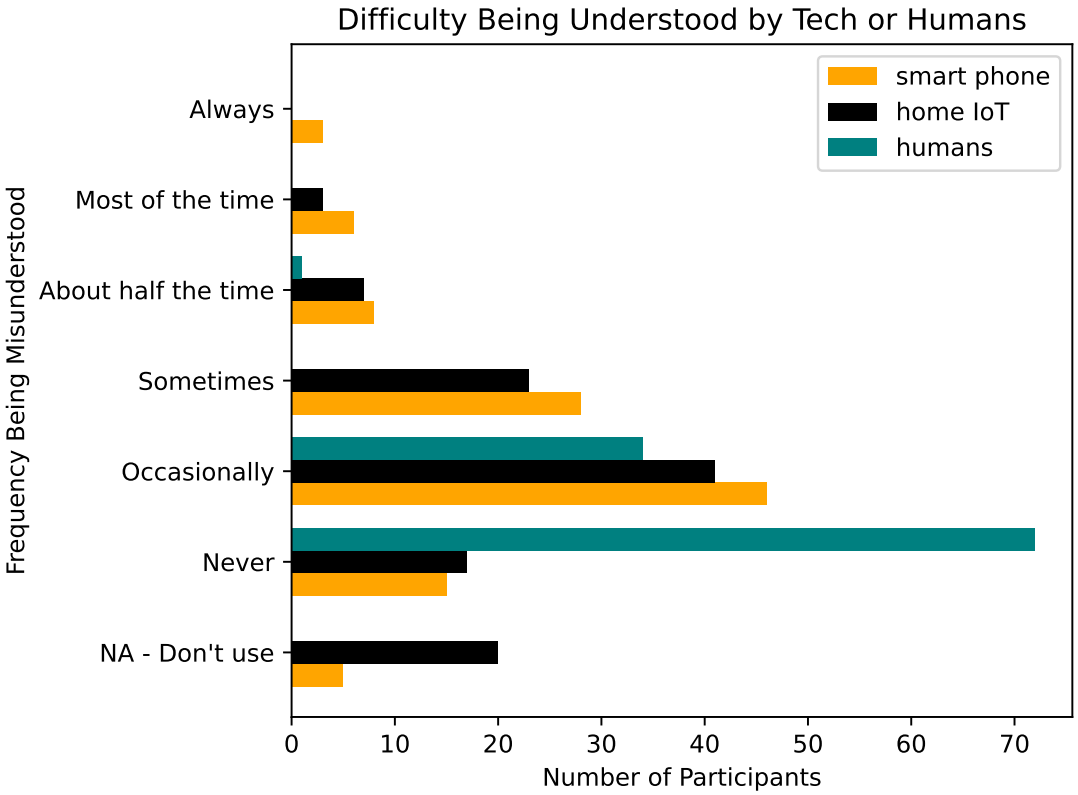


Figure B.2: Counts of participants rating frequency of having difficulty being understood by other English speakers, smart phones, or home IoT devices.

When asked how much they agreed with the statement, "Voice assistants usually understand me," 40 participants selected Strongly Agree, 54 Somewhat Agree, 11 Somewhat disagree, and one Strongly disagree. For the statement, "I am able to use voice assistants equally as well as the average or standard user they're developed for," 55 participants selected Strongly Agree, 33 Somewhat Agree, eight Somewhat disagree, and three Strongly disagree.

For the following statement, "It is important to me that speech recognition technologies, such as voice assistants, accurately understand my individual voice." 51 selected Strongly Agree, 42 Somewhat Agree, five Somewhat Disagree, and three Strongly Disagree. For the statement, "It is important to me that speech recognition technologies, such as voice assistants, accurately understand voices that sound similar to mine, for example, with the same accent and dialect," 53 participants selected Strongly Agree, 37 Somewhat Agree, seven Somewhat Disagree, and two Strongly Disagree.

| Age (Years) | | Gender Identity | | Household Income | |
|---|---|---|---|---|---|
| 21 to 30 | 25.2% | Female | 57.7% | Less than $50,000 | 50.5% |
| 31 to 40 | 30.6% | Male | 39.6% | $50,000 to $99,999 | 36.9% |
| 41 to 50 | 23.4% | Non-binary | 2.7% | $100,000 to $149,999 | 9.0% |
| 51 to 60 | 18.0% | Transgender | 1.8% | Above $150,000 | 2.7% |
| 61 to 70 | 2.7% | | | No response | 0.9% |

| Ethnic or Racial Identity | | Dialect Group | | Education | |
|---|---|---|---|---|---|
| White | 51.4% | SAE-Only | 22.5% | Less than high school | 1.8% |
| Black or African American | 48.6% | APP-Only | 18.0% | Graduated high school | 13.5% |
| Asian | 0.9% | AAE-Only | 7.2% | Some college education | 27.0% |
| American Indian or Alaska Native | 0.9% | Southern-Only | 21.6% | Associate's degree | 14.4% |
| | | SAE-Appalachian | 9.0% | Bachelor's degree | 32.4% |
| | | SAE-AAE | 21.6% | Master's degree | 9.9% |
| | | | | Trade school diploma | 0.9% |

| State (Current Residence) | | State of Origin | |
|---|---|---|---|
| Alabama | 1.8% | Alabama | 1.8% |
| Arkansas | 1.8% | Arkansas | 1.8% |
| Georgia | 8.1% | Georgia | 7.2% |
| Kentucky | 25.2% | Kentucky | 23.4% |
| Louisiana | 9.9% | Louisiana | 11.7% |
| Mississippi | 0.9% | Maryland | 0.9% |
| North Carolina | 9.9% | Missouri | 0.9% |
| South Carolina | 2.7% | North Carolina | 6.3% |
| Tennessee | 19.8% | South Carolina | 2.7% |
| Virginia | 9.0% | Tennessee | 21.6% |
| West Virginia | 10.8% | Texas | 0.9% |
| Virginia | 9.9% | | |
| West Virginia | 10.8% | | |

Table B.2: Summary of participant demographics.

| Recruitment Dialect Group | Criteria Based on Self-reported Presentation | No. |
|---|---|---|
| African American English (AAE) | speak AAE often/always (may also speak Southern often/always) AND not SAE often/always | 8 |
| Appalachian English (Appalachian) | speak Appalachian often/always (may also speak SAE or Southern often/always) AND not SAE often/always | 20 |
| Southern English only (Southern-only) | speak only Southern often/always (not AAE, Appalachian, or SAE) | 24 |
| Standardized American English and AAE (SAE-AAE) | speak AAE and SAE often/always (can include Southern) | 24 |
| Standardized American English and Appalachian (SAE-Appalachian) | speak Appalachian and SAE often/always (can include Southern) | 10 |
| Standardized American English only (SAE-only) | speak only SAE often/always (not AAE, Appalachian, or Southern) | 25 |

Table B.3: Dialect groups we recruited as part of our purposive sampling method, criteria, and number of participants recruited per group. Criteria are based on combinations of self-reported dialects spoken: African American English (AAE), Appalachian English (Appalachian), Standardized American English (SAE), and Southern English (Southern).

## B.4 Charts of Acceptability Results

We provide visual representation of the acceptability Likert ratings presented in Section 3.6.2.

### B.4.1 Acceptability of AI or "Software"

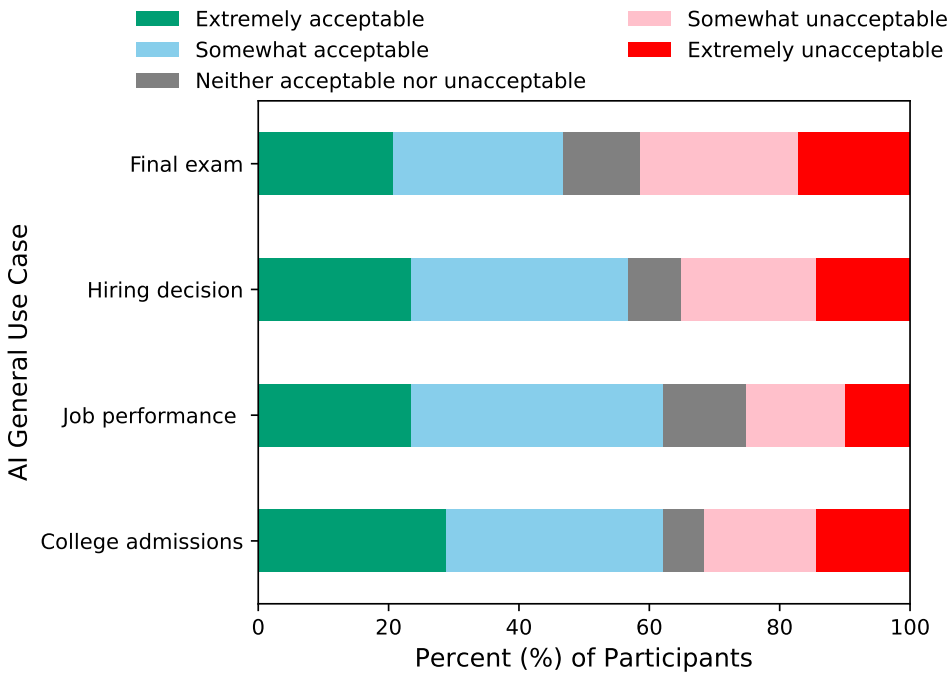


Figure B.3: Acceptability of the use of "software" for decision making in four use cases: college admissions interviews, final exam grading, job performance evaluation, and hiring interviews.

### B.4.2 Relative Acceptability of AI-enabled Voice Analysis

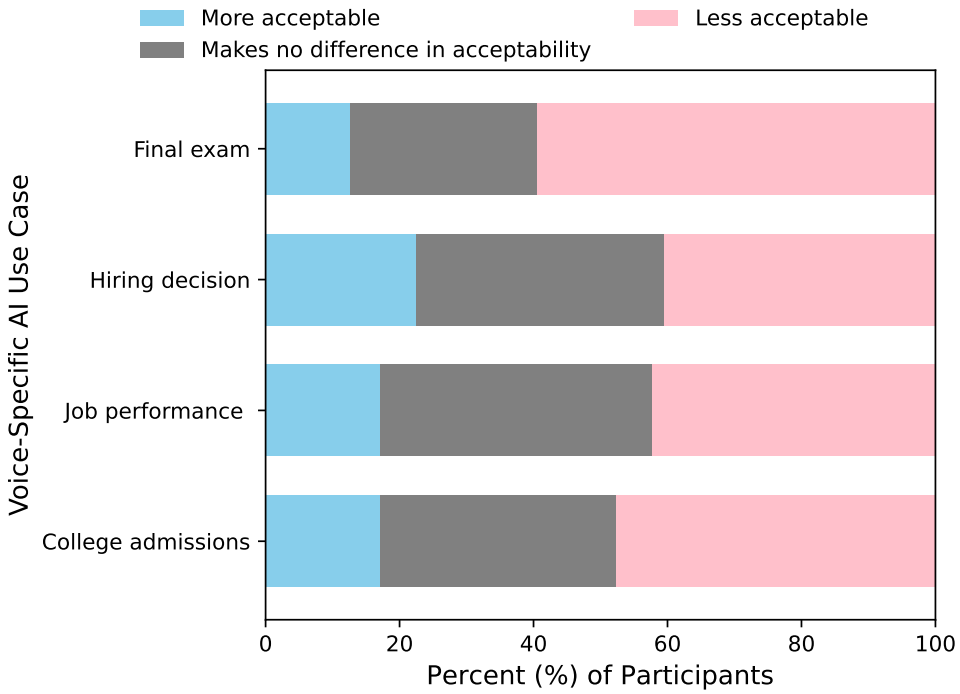


Figure B.4: Change in acceptability when we specified that AI-enabled analysis of voice data would be used for decision making in four use cases: college admissions, final exam grading, job performance evaluation, and hiring decision.

# Appendix C

# Interdisciplinary Approaches to Cybervulnerability Impact Assessment for Energy Critical Infrastructure (2024)

## C.1 Positive and Negative Perceptions of SME Groups

Below we present the positive and negative perceptions of both expert groups. Since the distribution of perceptions of both SME groups were very similar, as shown in Appendix C.2, we present results about the perception of SME groups in aggregate, not dividing perceptions according to the SME group to which the participant belonged but rather focusing on the target group of the participant's statements.

When discussing differences in the groups' approaches, more participants spoke positively about cyber SMEs than about energy OT SMEs for topics relating to Accessibility, Attack, and Vulnerability, but cyber SMEs were perceived more negatively for the topics of Consequence and Connectivity. Energy OT SMEs were depicted more positively regarding Connectivity and Consequence topics, and more negatively for Accessibility themes. The term "positive" means that the described group was perceived as adequately or effectively considering the factors relating to the topic, and "negative" signifies the converse.

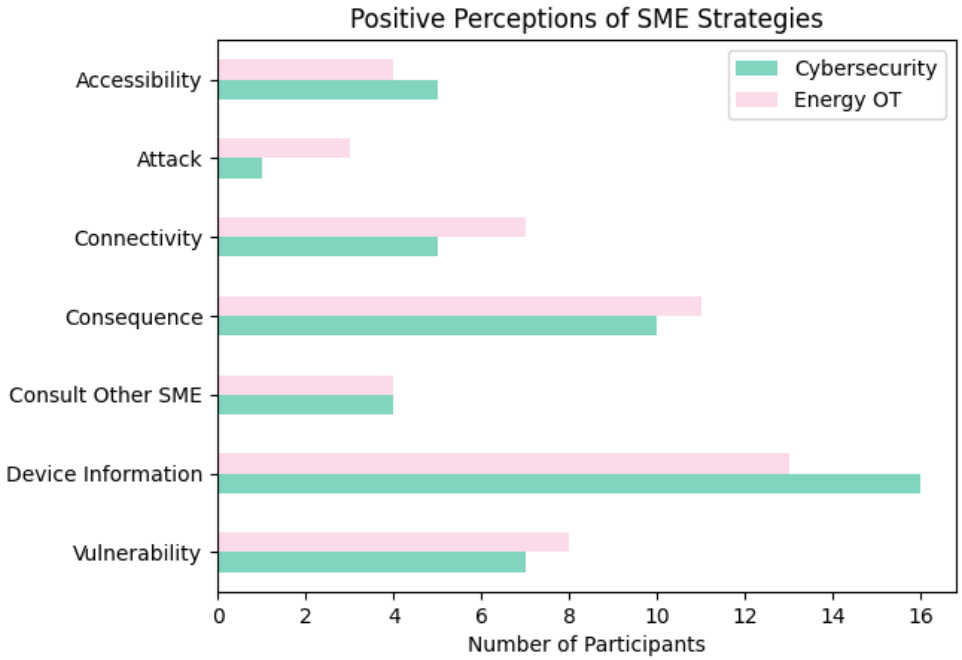


Figure C.1: Positive perceptions of each SME group's impact assessment strategies and understanding, using top-level strategy codes and showing count per participant by target group.

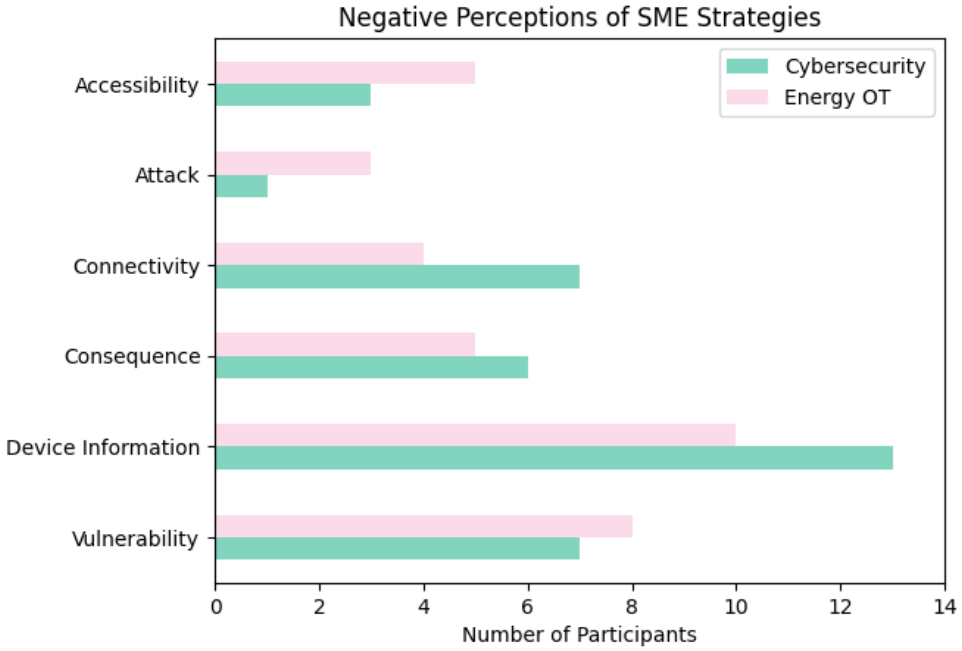


Figure C.2: Negative perceptions of each SME group's impact assessment strategies and understanding, using top-level strategy codes and showing count per participant by target group.

## C.2 Perceptions By Perceiving Group

The positive and negative perceptions offered by both SME groups about themselves and the other group were similar across each group, as can be seen in Table C.1 and Table C.2 below.

| Cyber Main Code | Positive | | Negative | |
|---|---|---|---|---|
| | Cyber | OT | Cyber | OT |
| Accessibility | 4 | 1 | 1 | 2 |
| Attack | 1 | 0 | 1 | 0 |
| Connectivity | 2 | 3 | 3 | 4 |
| Consequence | 4 | 6 | 5 | 1 |
| Consult Other SME | 2 | 2 | 0 | 0 |
| Device Information | 8 | 8 | 7 | 6 |
| Vulnerability | 6 | 1 | 1 | 6 |
| Total | 27 | 22 | 18 | 19 |

Table C.1: Positive and negative valences for Cyber SMEs' narrated depictions of Cyber and Energy OT SMEs' vulnerability impact assessment approaches, using our top-level codes for impact assessment topics.

| OT Main Code | Positive | | Negative | |
|---|---|---|---|---|
| | Cyber | OT | Cyber | OT |
| Accessibility | 3 | 1 | 2 | 3 |
| Attack | 3 | 0 | 1 | 2 |
| Connectivity | 4 | 3 | 1 | 3 |
| Consequence | 4 | 7 | 4 | 1 |
| Consult Other SME | 2 | 2 | 0 | 0 |
| Device Information | 5 | 8 | 8 | 8 |
| Vulnerability | 5 | 3 | 2 | 6 |
| Total | 23 | 21 | 16 | 19 |

Table C.2: Positive and negative valences for Energy OT SMEs' narrated depictions of Cyber and Energy OT SMEs, using our top-level codes for impact assessment topics.

## C.3 Interview Questions

Below are the interview questions we used in our semi-structured interviews.

### C.3.1 Background Questions

1. What is your current job role?
2. How many years?
3. Previous relevant experience?
4. What do you mainly work on? What is your broad area of expertise, for example, electric, ONG, computer security, or other?
5. How familiar are you with energy sector operational technology (OT)?
6. Do you have a background in or exposure to cybersecurity? *(If yes: In vulnerability analysis?)*
7. Have you done impact assessments?
8. Has your work focused on the impact of cyber vulnerabilities?
9. Have you encountered or do you know of any standard procedures, strategies, or metrics for assessing the impact of cyber vulnerabilities?
10. Have you worked on the same team as [energy OT/cybersecurity SMEs] before?
11. Has your work overlapped with their work?
12. In the context of OT security, have you collaborated across operational security and computer IT security teams?

### C.3.2 Self-Reported Strategies

1. How would you go about considering the potential impact of a cyber vulnerability in an OT system?
   (a) What information would you seek?
   (b) What questions would you ask?
   (c) How might the sub-sector in which the system is used influence your considerations, for example, generation, transmission, distribution?
   (d) How might the context or use-case of the system influence your approach?
   (e) How might the vendor of the system influence your approach?
   (f) Are there any other factors you would need or want to know about, to determine the impact, which we haven't already discussed?
2. In your experience, how do energy OT SMEs' approaches or methods in assessing impact of cybervulnerabilites differ from cybersecurity SMEs' approaches?
3. In your experience, how do energy OT SMEs' understanding of cyber vulnerabilities differ from how cybersecurity SMEs might understand cyber vulnerabilities?

### C.3.3 Perceptions of SME Groups

1. What are some gaps or differences in thinking or strategy that you have noticed in cybersecurity SMEs and energy OT SMEs that might be important to consider when conducting an impact assessment for an energy OT device or component?
2. What are some specific things [energy OT/cybersecurity] SMEs are neglecting or perhaps overprioritizing at the expense of other things?
   (a) *(Ask for other SME type.)*
3. What type of information needs to cross between the two domains to more accurately gauge cyberattacks' impact on energy OT systems?
4. Based on your knowledge and experience, what cyber vulnerability impacts might energy OT SMEs tend to underestimate or overestimate?
5. What cyber vulnerability impacts might cybersecurity SMEs tend to underestimate or overestimate?

## C.4 Appendix - Code Books

We include two code books: 1) vulnerability impact assessment strategy codes we used to label responses to all questions analyzed, in Table C.3, and 2) the perception codes used to label responses about participants' perceptions of general skills, motivations, and stereotypes of both expert groups, in Table C.4.

### C.4.1 Vulnerability Impact Assessment Strategy Codes

| Main Code | Subcode | Definition |
|---|---|---|
| Accessibility | General | What type of access has to be necessary to get to it? |
| | Network Access | Can the network be reached? Does an attacker have to have a local presence on a network, behind several firewalls; via internet? |
| | Physical Access | Does reaching the system require physical access or contact with the system? |
| Attack | General | Consideration of adversarial threat, attacker and potential attack vectors |
| Connectivity | Communication Protocol | Specific protocol configuration and modules |
| | Logical Connectivity | How does this relate to, influence, or control other things in the system, beyond just physical or network connection? Follow-on effects. Dependencies. |
| | Network Architecture | Overall map that defines network on large scale |
| | Network Connectivity | How does it interact on the network? What connections are open? |
| | Physical Connectivity | What are the physical devices, ports, wires, etc. that it is connected to? |
| | Segmentation | How segmented is it from networks or the outside world? Boundaries. |
| | Cost/Financial Impact | Impact on costs, finances, business priorities |
| | Damage to Equipment | Physical damage to device/system/equipment |

*Continued on next page*

*Table continued from previous page*

| Main Code | Subcode | Definition |
|---|---|---|
| Consequence | Disrupt Operations | Disrupt operations, shut down power, cause outage |
| | Ecological Impact | Effect on the environment |
| | General | Consequences, implications, impacts, what could happen; includes concepts like data loss, criticality, critical infrastructure, severity of impact |
| | Human Impact | How many people might be impacted? Could there be injuries? |
| | Remediation | What will it take to restore service? How long will it take to fix things? |
| | Scale of Impact | How far does the impact spread, e.g., region, duration, number of systems |
| Consult other SME | General | Need or wish to consult the other kind of SME to seek their explanation, advice, or work on part of the problem |
| Device Information | Basic Information | What is the device or system? How is it installed, configured? |
| | Capabilities | What can be done on this device? Can it be used beyond intended function? |
| | Functional Purpose | What does the device or system do? What is it made to do? How is it used? |
| | Maintenance History | Details about when, how, how often the system is updated or maintained |
| | Physical Location | Where does it live? Physical location of device or system. |
| | Protections | How is the system protected? Includes physical and network protections |
| | System Architecture | Information about setup in relation to other things in a broader system |
| | Ubiquity/Deployment | How common or prevalent is this device in the system and generally? |
| | Users | Who uses the system or device? |
| Vendor | General | Unprompted mention of the vendor, before being prompted directly |
| Vulnerability | Affected Area | What places, systems, or devices have the vulnerability? |
| | Exploit Requirements | What is required to exploit the vulnerability? How long would it take? |
| | Mitigation | Can the vulnerability be mitigated: prevented, stopped, or patched? |
| | Residual Impact | An impact that causes adverse effects that remain after efforts to remediate |
| | Understanding of Vulnerability | How does the vulnerability work? What is the vulnerability supposed to do? How does it relate to system operation? |
| | Vulnerability Information | Basic information about the vulnerability, e.g., whether it occurs or not, CVE number, vulnerability score, what system it applies to |

Table C.3: Thematic codes developed for responses to questions about participants' self-reported strategies for vulnerability impact assessment.

## C.4.2 Perceptions, Stereotypes, and Suggestions Codes

| Main Code | Subcode | Definition |
|---|---|---|
| Occupational Motivation (Cyber) | Protect computer systems | Cyber SME's goal is to protect computer systems, prevent intrusion |
| | Identify attack/exploit | Cyber SME seeks and identifies potential attacks or exploits |
| | Identify flaws or problems | Cyber SME seeks and identifies exploitable flaws or problems in software or hardware |
| | Identify things others haven't seen yet | Cyber SME identifies overlooked aspects of technology that could be exploited |
| | Identify vulnerabilities | Cyber SME identifies exploitable vulnerabilities |
| | Tear apart systems | Cyber SME takes apart systems to better understand them |
| Occupational Motivation (Energy OT) | Development/design | OT focuses on and prioritizes development or design |
| | Make sure the system works | OT focuses on and prioritizes maintenance, operations, making sure devices work and power is flowing |
| | Operational efficiency | OT focuses on and prioritizes reliability, minimizing or saving costs, saving time, operating at peak efficiencies |
| | Protect systems | OT focuses on protecting systems and tools |
| | Safety | OT focuses on and prioritizes human safety |
| Stereotypes | Cyber cuts off access to protect system | Cyber SMEs place protections on the system that prevent or make it difficult for OT to access or operate systems |
| | Cyber is detail-oriented | Cyber SMEs pay attention to details, go into rabbit holes |
| | Cyber lacks funding | There is a lack funding or resources to support cyber SMEs |
| | Cyber lacks understanding | Cyber SMEs don't understand some aspect |
| | Cyber overemphasizes | Cyber SMEs focus too much on some aspect |
| | Cyber overestimates | Cyber SMEs miscalculate or incorrectly consider some aspect as more severe, or consequential than it is |
| | Cyber sees computers | For Cyber SMEs, OT devices can be reduced to computers |
| | Cyber underestimates | Cyber SMEs do not sufficiently consider some aspect |
| | Cyber understands | Cyber SMEs are knowledgable and skilled in some aspect |
| | OT conflates security and safety | OT SMEs mistake safety or reliability for security of computer systems |
| | OT lacks funding | There is a lack funding or resources to support OT SMEs |
| | OT lacks imagination | OT SMEs do not, cannot, or find it difficult to think of possibilities outside of what they already know |
| | OT lacks understanding | OT SMEs don't understand some aspect |
| | OT overemphasizes | OT SMEs focus too much on some aspect |
| | OT overestimates | OT SMEs miscalculate or incorrectly consider some aspect as more severe or consequential than it is |
| | OT takes shortcuts | OT SMEs leave access open, create backdoors, or otherwise leave open vulnerabilities for the sake of having easy access |
| | OT underestimates | OT SMEs do not sufficiently consider some aspect |
| | OT understands | OT SMEs are knowledgable and skilled in some aspect |
| Suggestion | Collaboration | SME suggests how SMEs should or can collaborate |
| | Make systems usable for OT | SME suggests that security design or policies should also ensure that systems can still be accessed and used by OT |
| | Teach Cyber | SME suggests that Cyber should know or learn something |
| | Teach OT | SME suggests that OT should know or learn something |

*Continued on next page*

*Table continued from previous page*

| Main Code | Subcode | Definition |
|---|---|---|

Table C.4: Thematic codes responses to questions about participants' perceptions of the two SME groups' strategies for vulnerability impact assessment and understanding of vulnerabilities.

# C.5 Subcodes

Below are counts per participant for the subcodes described in Appendix C.4.

## C.5.1 Self-reported impact assessment strategies

| Main Code | Subcode | Cybersecurity | Energy OT | Total Count |
|---|---|---|---|---|
| Accessibility | General | 6 | 6 | 12 |
| | Network Access | 5 | 1 | 6 |
| | Physical Access | 4 | 2 | 6 |
| Attack | General | 4 | 5 | 9 |
| Connectivity | Communication Protocol | 1 | 1 | 2 |
| | Logical Connectivity | 7 | 10 | 17 |
| | Network Architecture | 2 | 1 | 3 |
| | Network Connectivity | 4 | 3 | 7 |
| | Physical Connectivity | 1 | 1 | 2 |
| | Segmentation | 1 | 2 | 3 |
| Consequence | Cost or Financial Impact | 2 | 3 | 5 |
| | Damage to equipment | 2 | 1 | 3 |
| | Disrupt Operations | 2 | 4 | 6 |
| | Ecological impact | 1 | 0 | 1 |
| | General | 7 | 8 | 15 |
| | Human Impact | 5 | 2 | 7 |
| | Remediation | 1 | 4 | 5 |
| | Scale of Impact | 6 | 5 | 11 |
| Consult other SME | General | 3 | 3 | 6 |
| Device information | Basic Info | 5 | 7 | 12 |
| | Capabilities | 3 | 3 | 6 |
| | Functional Purpose | 4 | 5 | 9 |
| | Maintenance | 0 | 1 | 1 |
| | Physical Location | 2 | 3 | 5 |
| | Protections | 5 | 4 | 9 |
| | System Architecture | 4 | 5 | 9 |
| | Type of Facility | 0 | 1 | 1 |
| | Ubiquity/Deployment | 2 | 2 | 4 |
| | Users | 1 | 1 | 2 |
| Vendor | General | 3 | 2 | 5 |
| Vulnerability | Affected area | 1 | 2 | 3 |
| | Exploit Technical Requirements | 5 | 3 | 8 |
| | Mitigation | 1 | 4 | 5 |
| | Residual Impact | 1 | 3 | 4 |
| | Understanding of Vulnerability | 7 | 4 | 11 |
| | Vulnerability Information | 6 | 4 | 10 |

Table C.5: Strategy subcodes applied to each individual's self-reported impact assessment strategies and considerations, showing count per narrating participant based on their expert group, and also showing total count.

## C.5.2 Counts of perceptions of the SME groups

Table C.6 includes all codes for perceptions of the SME groups, counted per participant (from either group) who expressed a negative opinion about cyber SMEs ("Cyber Negative"), a positive opinion about cyber SMEs ("Cyber Positive"), a negative opinion about energy OT SMEs ("OT Negative"), or a positive opinion about energy OT SMEs("OT Positive").

| Main Code | Subcode | Cyber Negative | Cyber Positive | OT Negative | OT Positive |
|---|---|---|---|---|---|
| Accessibility | General | 3 | 7 | 3 | 2 |
| | Network Access | 0 | 0 | 2 | 0 |
| | Physical Access | 0 | 0 | 3 | 1 |
| Attack | General | 2 | 4 | 2 | 0 |
| Connectivity | Communication Protocol | 1 | 1 | 0 | 0 |
| | Logical Connectivity | 4 | 5 | 4 | 5 |
| | Network Architecture | 0 | 0 | 1 | 1 |
| | Network Connectivity | 0 | 1 | 1 | 1 |
| | Physical Connectivity | 0 | 0 | 2 | 0 |
| Consequence | Damage to Equipment | 1 | 0 | 0 | 1 |
| | Disrupt Operations | 1 | 0 | 0 | 2 |
| | General | 7 | 6 | 2 | 12 |
| | Human Impact | 0 | 0 | 0 | 1 |
| | Scale | 0 | 2 | 0 | 2 |
| Consult Other SME | General | 0 | 4 | 0 | 4 |
| Device Information | Basic Info | 1 | 1 | 1 | 4 |
| | Capabilities | 3 | 11 | 9 | 0 |
| | Functional Purpose | 5 | 4 | 1 | 12 |
| | Maintenance History | 1 | 0 | 0 | 1 |
| | Physical Location | 0 | 0 | 0 | 1 |
| | Protections | 2 | 0 | 1 | 0 |
| | System Architecture | 4 | 1 | 0 | 7 |
| | Ubiquity/Deployment | 1 | 0 | 0 | 3 |
| Vulnerability | Exploit Technical Requirements | 3 | 1 | 1 | 0 |
| | General | 0 | 3 | 3 | 1 |
| | Mitigation | 1 | 0 | 0 | 1 |
| | Residual Impact | 0 | 1 | 0 | 0 |
| | Understanding of Vulnerability | 0 | 9 | 10 | 0 |
| | Vulnerability Information | 0 | 2 | 1 | 2 |

Table C.6: Strategy subcodes applied to participants' stated perceptions of the SME groups' strategies and understanding, showing counts per SME group being characterized (target of the comment).

# Appendix D

# Translation and AI Health Assistants in Healthcare Contexts

## D.1 Screening Survey

We include the screening survey questions below in English. Screening surveys were distributed in two languages, Mandarin and Spanish. Note that the question "Do you speak standard Mandarin?" was only asked in the Mandarin-language survey. Where both languages are listed below, in the actual survey we only included the language of the survey. For our peer-reviewed paper submission, we will make the Spanish and Mandarin versions of the screening survey available to readers.

### Contact Information and Interview Modality

1. **How can we send you information about the interview?**
    - Email
    - Phone call (through a US phone company)
    - Text message
    - WhatsApp
    - WeChat
2. **Do you prefer a Zoom interview (on the computer) or an in-person interview?**

### Language

3. **What is your primary language?** "Primary language" refers to the language that you use most often or that you are most comfortable speaking every day.
    - Mandarin/Spanish
    - English
    - Other:
4. **Do you speak standard Mandarin?** (普通话/国语/华语吗)
5. **How well do you speak English?**
    - Very well
    - Well
    - Not well
    - I do not speak English / I can't speak it at all

- Prefer not to respond

6. **How well do you read English?**
   - Very well
   - Well
   - Not well
   - I do not read English / I can't read it at all
   - Prefer not to respond

## Medical Visits

In the questions below, when we say "doctor" we mean any medical provider you have seen or medical center that you have visited. For example: a nurse, dentist, psychologist, physical therapist, or a medical clinic, hospital, or pharmacy. When answering these questions, focus solely on visits for your own medical care (not for other members of your family).

7. **In the last year, how many times have you visited a doctor?**
   - Never
   - 1-5 times
   - 6-10 times
   - 11-15 times
   - Over 15 times
   - I don't remember
8. **In what city do you usually visit the doctor?**
9. **After seeing the doctor, have you ever received written medical advice in English?**
10. **Have you ever tried to read medication instructions in English?**

## Interpretation

Interpretation is when someone helps you understand another person who is speaking a different language. An interpreter can help you understand the English words the doctor says by telling you the same thing in Spanish/Mandarin. This can be in person or over the phone.

11. **Which of the following people or apps helped you understand doctors and medical staff when they spoke in English?** Check all that apply.
    - A professional interpreter
    - A family member or friend
    - A translation app on a phone or tablet
    - Other:
    - I didn't receive any help

## Demographics

We ask the following questions to learn about the different kinds of people who are taking part in our survey.

12. **What age group are you in?**
    - 18-24
    - 25-34
    - 35-44

- 45-54
- 55-64
- 56-60
- 65-74
- 75-84
- 85+
- Prefer not to respond

13. **How do you describe your gender? (You can choose more than one.)**
    - Man or male
    - Woman or female
    - Non-binary
    - Transgender
    - Prefer to self describe:
    - Prefer not to respond
14. **What is the highest level of school that you have completed? (If you attended school in another country, choose the level that is the same.)**
    - Primary school
    - Secondary school (high school diploma or equivalent including GED)
    - Trade school diploma, certificate, or license
    - Apprenticeship (for example, as a chef or carpenter)
    - Some college but no degree
    - Associate degree in college (2-year)
    - Bachelor's degree in college (4-year)
    - Master's degree
    - Doctoral degree
    - Professional degree (JD, MBA, MD)
    - Other:
    - Prefer not to respond
15. **How much money did your family make in 2024 before taxes?**
    - $0 - $29,999
    - $30,000 - $59,999
    - $60,000 - $89,999
    - $90,000 - $119,999
    - $120,000 - $149,999
    - $150,000 or more
    - Prefer not to respond
16. **Do you have a degree or certificate in computer science?**
17. **Do you have a degree or certificate in medicine, nursing, or pharmacology?**

## D.2 Interview Questions

The following are the interview questions used in the study, presented in English. Given the semi-structured nature of the interview, we did not ask every question in every interview.

### Experience with Language Services at Doctor's Office

1. In prior experiences going to the doctor, what actions have you taken to communicate with doctors who only speak English?
   (a) Which of those options do you use the most?
   (b) What app did you use?
   (c) Who helped you?
2. Could you tell us in a sentence or two: What is the most useful aspect? What works well?
3. What is the least satisfying aspect? What doesn't work well?
4. To achieve the best communication results, do you make preparations before consulting a doctor? For example, translate and write down some medical terms before the appointment.

### Interpreter Experience

5. Have you ever requested an interpreter at the doctor's office, that is, in person?
   (a) Could you tell us in a sentence or two: What is the most useful aspect? What works well?
   (b) Could you tell us in a sentence or two: What is the least satisfying aspect? What doesn't work well?
6. When you had an interpreter, was it in person or online (by phone or tablet)?
   (a) What do you think about the difference between in-person and phone/tablet contact?
   (b) Was it the same person every time?
7. Have you ever felt that the interpreter's translation was not correct or incomplete?
8. Have the interpreters come from the same cultural background as you, so they can fully grasp the meanings and implications of your words?
9. On a scale of 1 to 5, where 1 is "very unlikely" and 5 is "very likely": how likely would you be to allow an interpreter to interpret what you and the doctor were saying? Why?
10. On a scale of 1 to 5, how much do you trust an interpreter to translate your medical information? Why?
11. Imagine that you need to be hospitalized for several weeks and you were assigned an in-person interpreter.
    (a) How would you feel about having informal conversations with the interpreter, for example about your hobbies?
    (b) How likely would you be to allow the interpreter to mention to the doctor something relevant to your health that he or she had learned about you in an informal conversation while interpreting?
12. What are your privacy concerns when using an interpreter?

### Translation App Experience

13. Have you ever used a translation app?
    (a) What app did you use?
    (b) If no: Have you ever seen someone else use this type of app?
14. When using a translation app, have you ever felt like the translation was not correct or incomplete?
    (a) How do you know if the translation is not accurate or correct?
15. Have you ever been unable to use a translation app because you had no reception?

16. On a scale of 1 to 5, how likely would you be to use a translation app to interpret what you and the doctor are saying? Why?
17. On a scale of 1 to 5, how much do you trust a translation app to translate your medical information? Why?

## Scenario (English-only Instructions)

18. Given the situation that I just described, what would you do first to try to understand and follow the instructions?
19. Given what you just said that you would do, is this what you have done in the past in similar circumstances?
20. Have you received documents/instructions exclusively in English from the clinic?
    (a) Have you ever asked for a written translation at the doctor's office?
    (b) Have you ever received an unsolicited translation?
    (c) And from the pharmacy, have you received documents/instructions exclusively in English?
21. What have you done to understand words you don't know in English?
22. What do you think about using a translation app to translate the medical instructions in this scenario?
23. What do you think about asking someone you know, like a friend or family member, to help?
    (a) Have you ever asked someone to help you with something like this?
24. On a scale of 1 to 5, how much do you trust a friend or family member to translate your medical information? Why?
25. On a scale of 1 to 5, how likely would you be to ask a friend or family member for help translating the instructions? Why?

## Artificial Intelligence Health Assistant

26. Have you ever used an AI chatbot like ChatGPT?
    (a) If no: Have you ever seen someone else use this type of app?
    (b) If no: Would you like me to explain what it is?
27. Have you ever used a voice assistant like Alexa, Siri, or Google Home?
    (a) If no: Have you ever seen someone else use this type of app?
    (b) If no: Would you like me to explain what it is?
28. What is the first thing that comes to mind, whether positive or negative, when you think about this imaginary app?
    (a) Can you think of one advantage or positive aspect?
    (b) Can you think of one disadvantage or negative aspect?
29. Would you ask the app to explain something you don't understand? What would you use this app for?

### AI Real-time Interpretation

30. In the medical clinic, for interpretation: would you prefer this artificial intelligence application or a professional human interpreter? Why?
31. Have you ever seen an English speaking doctor from home via a remote video conference?
    (a) If yes: was there an interpreter? How were you able to understand them?
    (b) Would it help you to see the English words the doctor is saying, like captions in a film (that is, a transcription)?
32. If the AI could translate/interpret for you during remote calls, would that change how often you have remote visits (versus going in person)?

### AI Digital Assistance: Contact office from home

33. Have you ever been at home and needed to contact the clinic, for example, to schedule a new appointment, or to request a medical document or bill?
    (a) If you could schedule an appointment or order documents using the AI assistant, speaking or writing to it in Spanish, and the app would act as your agent. Would you use this AI app? How would things change for you if the app could help you with things like scheduling appointments or ordering documents?

### Informal Chats with the AI Chatbot

34. How would you feel about having informal chats with the app?
    (a) Would you like this AI app to remember some of your personal details, such as your hobbies? Why or why not?
35. How likely would you be to allow the app to mention something relevant to your health to your doctor that it learned about you in a casual conversation?
36. Would you use the app for conversations other than translation? (For example, seeking emotional support, medical advice, or dietary recommendations during recovery, if participants don't understand.)

### AI Understanding and Using Dialects

37. Have you had any experience with a translation where the interpreter or application didn't translate well because of this variation?
38. If the app could change its accent and vocabulary depending on the user's country or region, how would you feel? Why? Would you like to program it to use your preferred regional vocabulary or accent?
39. Can you think of some examples of words or phrases that vary in different places and that an AI application would need to know in multiple ways?
40. What Spanish dialects or accents would you be surprised if the app could understand perfectly?
41. And what Latin American languages, for example, indigenous languages, would you be surprised if the app could fully understand?
42. Have you or your translation app ever had difficulty understanding a doctor because of their accent?

### AI Voice Cloning

43. When interpreting, the app might modify its voice to sound more like yours. How would you feel if the app had a voice almost identical or identical to yours?
    (a) And if you were talking to the doctor, would you sound like you do when you talk to the doctor?

## Privacy Concerns

44. What are your concerns about privacy?
45. Are you worried about who else might see your conversations?
46. Would you mind if the information was stored in the cloud or on a server owned by a technology company, like Google?
47. If the app can sell your data to advertising companies, how likely would you be to use the app?

## Humans Pros/Cons

48. In your opinion, what positive and useful things do humans do that might make you trust them more than you would trust an app?

49. What unhelpful, frustrating, or even harmful things do humans do that might make you trust them less than you would trust an app?

### Preferences

50. If you could choose anyone or anything to be your interpreter or translator, who or what would you choose? Why?
51. What factors are most important in your choices? For example, translation speed, your trust in the translator, your privacy concerns, the translator's level of knowledge of English or medical terminology, and the doctor's advice on which translation option to choose.

### Unmet Needs, Suggestions, Final Thoughts

52. In what other medical settings outside of the clinic do you think there aren't enough translation options for medical situations?
53. How does your experience vary across different types of doctors or providers, including specialists such as dentists, therapists, physical therapy, or eye doctors?
54. What translation needs in medical situations aren't being met now?
55. What suggestions or advice do you have for a hospital that is building this healthcare assistant that translates between patients and doctors?
56. Is there anything you think is important to consider that we didn't discuss today?

## D.3 Quote Translations

Q1. **Original:**
**Translation:** There are iPads that have the option to make video or phone calls. There are also physical interpreters, and almost always what I see is the same: the interpreters are from somewhere else. I've had some who are from Peru or Mexico, and the patient is from Puerto Rico or Cuba, and they use different words. Or they start using medical terminology, not the patient, but the doctors. And the interpreter gets lost. And the translation they try to give is wrong. So, obviously, there's no way to detect it if you don't know both languages and the medical terminology. But for me, who knows a little bit about everything, it's something I've noticed, and that's why I think that [knowledge of medical terminology] would be the most relevant thing for me in this situation. (ES09)

Q2. **Original:** Uno a veces dice algo, y cuando el traductor lo explica, no es lo que uno dijo, o no está completo, que uno tiene que interferir y decir, mira, es que faltó decir esta información o faltó decir esto. (ES14)
**Translation:** sometimes you say something, and when the translator [interpreter] explains it, it's not what you said, or it's incomplete, so that you have to intervene and say, look, you failed to mention this information or you failed to say this.

Q3. **Original:** Tenía dos días en el hospital y esa noche me tocó estar con él. Y pues, desde que yo llegué. Estuvimos hablando en español y todo bien. Cuando a mí me dan reporte del paciente reporte de enfermera enfermera, me dicen que el paciente había estado muy desorientado, que no hacía caso y que estaba muy confundido. Y así, cuando yo hablé con el paciente, el paciente entendía todo perfectamente. Y entonces una de las enfermeras me enseñó a porque ella llevó la aplicación, esa traductora al cuarto, hicimos videollamada con alguien de otro país. Y entonces le empezó a hacer preguntas al paciente y el paciente contestaba de forma incoherente en el sentido de que no sé, le hacían una pregunta de sí o no. y él contestaba otra cosa, pero en realidad el paciente tenía, O sea, no escuchaba bien el ipad. Tenía un volumen muy bajito. y además usaban términos que el paciente no usaba. Entonces, como él no sabía lo que le estaban preguntando, contestaba otra cosa. y por eso todos pensaron que él estaba confundido, que estaba desorientado. que no sabía qué estaba pasando, pero en realidad, todo el tiempo fue un problema de comunicación. (ES09)

**Translation:** I think he'd been in the hospital for two days, and that night I had to be with him. And so, from the moment I arrived, we spoke in Spanish and everything was fine. When they gave me the patient report, the nurse's report, they told me that the patient had been very disoriented, that he wasn't paying attention, and that he was very confused. When I spoke with the patient, the patient understood everything perfectly. And then one of the nurses showed me why. She brought the app [with] the interpreter into the room. We made a video call with someone from another country. And then he started asking the patient questions, and the patient answered incoherently, in the sense that, I don't know, they were asking him a yes or no question, and he answered something else. But in reality, the patient couldn't hear the iPad well. The volume was very low. And they also used terms that the patient didn't use. So, since he didn't know what they were asking him, he answered something else. And that's why everyone thought he was confused, that he was disoriented, that he didn't know what was going on. But in reality, it was a communication problem the whole time. (ES09)

Q4. **Original:** Bueno, realmente nunca he podido llamar a coger una cita porque hablan solo. Bueno, la gran mayoría hablan siempre en inglés y cuando ofrecen el servicio de traductor en línea, es muy, muy demorado. A veces no se les escucha bien. el traductor. (ES15)
**Translation:** I have never really been able to call to make an appointment, because the vast majority always speak English, and when they offer online translation services, it is very very delayed. Sometimes you can't really hear the interpreter.

Q5. **Original:** Como cuando estás recibiendo una noticia difícil o un diagnóstico difícil. Yo lo viví y en esa ocasión estaba la traductora en persona y sintió y vio mi dolor cuando me estaban dando el diagnóstico. Y para mí fue muy importante que ella simplemente me mirara con con resiliencia o con compasión. Aunque no me dijo nada porque no tenía por qué decirme nada. Era un asistente. Pero puso su mano en mi hombro y me dijo: todo va a estar bien. Y créanme, hasta la fecha, después de 8 años, no se me olvida. Y pienso que si hubiera estado un traductor en el teléfono azul [remote interpreter], pues él no ve mi reacción, ¿no? O no ve lo que está pasando. Y no podía haberme dicho nada. Los doctores, sí, pues de cierta manera, no sé, en sus protocolos, pero nada que ver. O sea, ellos, lo siento es lo máximo que pueden decir, ¿no? Pero esta persona vio y sintió mi dolor, y yo me sentí bien. Y siento que el IA, nunca lo haría. (ES07)
**Translation:** Like when you're receiving difficult news or a difficult diagnosis. I experienced it, and on that occasion, the translator was there in person, and she felt and saw my pain when they were giving me the diagnosis. And it was very important to me that she simply looked at me with resilience or compassion. Although she didn't say anything to me, because she didn't have to tell me anything. She was an assistant. But she put her hand on my shoulder and told me: everything will be okay. And believe me, to this day, after 8 years, I haven't forgotten it. And I think if there had been a translator on, like, the blue phone [remote interpreter], well, he doesn't see my reaction, right? Or he doesn't see what's happening. And he couldn't have told me anything. The doctors, yes, in a certain way, in their protocols, but it doesn't even compare. I mean, the most they can say is "I'm sorry," right? But this person saw and felt my pain, and I felt good, and I feel like AI would never do that. No. (ES07)

Q6. **Original:** Cuando son conversaciones tan largas, si definitivamente es más fácil que esté un intérprete. Es muy dispendioso. Tú estás ahí con el celular. Esperar que la persona te diga tú, leer, borrar, cambiar el idioma porque tú tienes que cambiar la flecha. Y a veces no la cambias porque estás, pero con el tiempo apremia, entonces no la cambias y tienes que volver a reiniciar y volver a repetir el mensaje y esperar que él lo lea, volver a borrarlo, cambiar el mensaje para que él me hable. Entonces es muy, muy poco práctico en un momento de esos hacer todo ese proceso en el celular. (ES08)
**Translation:** When the conversations are so long, it's definitely easier to have an interpreter there. It's very time-consuming. You're there with your cell phone. Waiting for the person to tell you, read, delete, change the language because you have to change the arrow. And sometimes you don't change it because you're there, but time is running out, so you don't change it and you have to restart and repeat the message again and wait for them to read it, delete it again, change the message so they can talk to you. So it's very, very impractical in a moment like that, to do that whole process on your cell phone. (ES08)

Q7. **Original:** Tal vez las cosas se pueden sacar de contexto … sobre todo si hacemos comentarios sarcásticos o algo así. Es algo que, pues no hay como detectar de parte de la inteligencia artificial. Y entonces son cosas que, pues podrían traducirse de una forma errónea. (ES09)

**Translation:** Perhaps things can be taken out of context … especially if we make sarcastic comments or something like that. It's something that artificial intelligence can't detect. And so these are things that could be translated incorrectly. (ES09)

Q8. **Original:** Lo hice. Lo tengo muy presente porque uno de mis hijos tiene una condición, y su medicamento era. tenía que ser exacto. Si cada detalle a detalle era, tiene que ser exacto, Así es que yo o anotaba. Y aún así me llevaba nota, y y todo porque sabía que me iban a dar instrucciones del medicamento. Y cuando llegaba a casa, o sea, yo no me podía arriesgar porque era una dosis muy precisa y muy peligrosa. Sí, así es que yo era de que tenía que traducir renglón a renglón para no equivocarme con su dosis y con su horario, porque era de diario un medicamento que él tiene que tomar, y era muy delicado. Sí, Sí, lo he hecho. (ES07)
**Translation:** I did this. I have this very present in my mind because one of my children has a condition, and his medication had to be exact. Every single detail had to be exact, so I wrote it down. And I still kept a note and everything, because I knew they were going to give me instructions for the medication. And when I got home, I mean, I couldn't take the risk because it was a very precise and very dangerous dose. Yes, so I had to translate line by line so I wouldn't make a mistake with his dosage and his schedule, because it was a daily medication that he has to take, and it was very delicate. Yes, yes, I've done it. (ES07)

Q9. **Original:** Ya he pasado varias veces, porque … cuando uno está en el hospital, realmente a veces lo que te dicen, te queda como la mitad, aunque te lo digan en español. Yo soy muy de que tengo que tomar nota. Así es que pues en ese momento no llevo lápiz y papel, así es que yo digo, Ok, pero me aseguro que venga en el papel que me van a dar, las instrucciones. Y pues ya digo, pues lo poquito que entendí, sino, lo traduzco llegando a casa, y es lo que hago, lo traduzco con un traductor en Internet. (ES07)
**Translation:** I've been through it several times, because … when you're in the hospital, sometimes what they tell you, you're left with half of it, even if they tell you in Spanish. I'm very much a person who has to take notes. So, at that moment, I don't have a pencil and paper. So I say, Okay, but I'll make sure it's on the paper they're going to give me, the instructions. And as I said, [if I don't understand,] I translate it when I get home, and that's what I do, I translate it with an online translator. (ES07)

Q10. **Original:** Como somos seres humanos, tenemos estados de ánimo, ¿cierto? Y no todos los días nos levantamos igual. Aunque queramos, tenemos problemas. A veces he notado que algunos [intérpretes] son como groseros y uno tiene que entender que de pronto esa persona estará pasando por un mal momento. O sea, tiene que trabajar porque tiene que ganarse la vida, pero somos seres humanos, entonces tenemos estado de ánimo. Entonces muchas veces los estados de ánimo influyen en el servicio que el intérprete ofrece. (ES14)
**Translation:** Since we're human beings, we have moods, right? And we don't wake up the same way every day. Even if we want to, we still have problems. Sometimes I've noticed that some [interpreters] are kind of rude, and one has to understand that that person might be going through a bad time. I mean, they have to work because they have to earn a living, but we're human beings, so we have moods. So, moods often influence the service the interpreter provides. (ES14)

Q11. **Original:** piense que estoy, no sé, exagerando con lo que le digo, y simplemente él traduzca lo que le conviene, lo que él quiere, no exactamente, lo que le estoy diciendo. (ES02)
**Translation:** in the way they find convenient, not what I'm saying, just say[ing] whatever he/she wants

Q12. **Original:** Yo creo que estaría muy bien, pero me pongo también a ver, o sea, en la parte de la programación se podría complicar mucho y el hecho de que se complique podría causar errores. Entonces, pienso yo que, o sea, si, como se dice en el español hay muchos acentos y todo, pero a la hora de términos médicos o traducciones médicas, siento que todo es algo como neutral. No necesariamente necesitas un acento en específico para que las personas de habla hispana puedan entender. Entonces, creo yo, que con un o sea con solo el español normal, o sea promedio. Sí se podría llegar a entender toda la aplicación. (ES06)
**Translation:** I think it would be great, but I'm also thinking, I mean, the programming part could get very complicated, and the fact that it gets complicated could cause errors. So, I think that, I mean, yes, as they say in Spanish, there are a lot of accents and everything, but when it comes to medical terms or medical translations, I feel like everything is somewhat neutral. You don't necessarily need a specific accent for Spanish speakers to understand. So, I think that with just normal, average Spanish, you could understand the entire application. (ES06)

Q13. **Original:** Mira, si es igual a mi acento, yo me sentiría feliz porque es como si estuviera un consultorio en mi país y me entendería. "Doctor: me duele la cabeza, tengo un yeyo." Un yeyo que es síntoma de malestar, Estoy mareada. Y Si me contesta en mi acento, eso sería—yo creo que hasta me mejoraría, de la felicidad. (ES03)
**Translation:** Look, if it's the same as my accent, I'd be happy because it's like I'm in a doctor's office back home and they'd understand me: "Doctor: My head hurts, I have a yeyo." A yeyo, that's a sign of discomfort. "I'm dizzy." And if it answers me in my accent, that would be—I think I'd even feel better, from happiness. (ES03)

Q14. **Original:** Ahí sí sería diferente si fuera que yo mandara una nota de voz, y la aplicación hiciera como mi voz, pero para que el médico la escuchara como mi voz sería muy diferente. Sería muy bueno ahí. Sí, me gustaría. … Porque muchas veces uno reconoce a las personas por su voz. … Entonces pienso que ayuda a la comunicación. (ES14)
**Translation:** In that case, it would be different if I were to send a voice note, and the app would act like my voice, but for the doctor to hear it as my voice. That would be very different. That would be very good. Yes, I would like it. … Because often you recognize people by their voice. … So I think it helps communication. (ES14)

Q15. **Original:** Sí creo que sería más natural, pero al inicio, sí me parecería como raro o extraño, pero creo que justo no se pierde lo que decía anteriormente, porque pienso que esta aplicación básicamente eliminaría esa brecha que yo siento que pasa entre los intérpretes. (ES21)
**Translation:** Yes, I think it would be more natural, but at first it would seem strange or odd to me, but I then think it doesn't lose what I was saying before [sentiment], because I think this application would basically eliminate that gap that I feel exists between interpreters. (ES21)

Q16. **Original:** me he visto en la obligación de pedirle un familiar que que llame por mí, o sea, para para poder hacerlo porque no tengo, No, no, No tengo otro modo de de poder comunicarme con alguien para pedir una consulta, si no le puedo entender. (ES11)
**Translation:** I have been obligated to ask a family member to call for me, because I don't have another way to communicate with someone to ask for an appointment, if I can't understand them.

Q17. **Original:** El solo hecho de pedir una cita médica … es complicado a veces, si me entiendes, por el tema del idioma. Si tengo una aplicación que me puede ayudar a pedir una cita médica, sería buenísimo. (ES14)
**Translation:** Just the act of making a medical appointment … is complicated sometimes, due to the language issue. If I have an app that can help me ask for a medical appointment, that would be great. (ES14)

Q18. **Original:** Porque si yo llamo a la clínica, yo no entendería lo que ellos me dijeran porque como ellos hablan inglés. (ES03)
**Translation:** Because if I call the clinic, I wouldn't understand what they were saying, since they speak English.

Q19. **Original:** Si vas a una emergencia con un niño. Y si no tienen traductor, tienen esa computadora para traducir. Realmente uno como mamá tiene que ver la forma de que te atienden al hijo, no porque vas a dejar—que no sabes leer el inglés y no pides ayuda. Tu hijo te está muriendo. ¿Qué puede hacer uno? Es mejor decir, no sé en inglés, ¿Puede ponerme un traductor, por favor? Porque urgen ver que el niño, ¿qué está pasando? Y ahí, si te está traduciendo una persona o un aparato o un teléfono, lo que quieres ver tú como mamá: que esté bien tu niño. [...] Porque ver a tu hijo enfermo, y no sabes lo que está pasando. Tú quieres que te lo atiende. (ES04)
**Translation:** If you go to an emergency room with a child. And if they don't have an interpreter, they have that computer to translate. As a mother, it really has to do with how they treat your child, because you're not just going to leave it alone, because you don't know how to read English and don't ask for help. Your child is dying on you. What can you do? It's better to say, I don't speak English, can you get me a translator, please? Because they urgently need to see the child, What is happening? And then, if a person or a device or a phone is translating for you, what you want to see as a mother: that your child is okay. Because seeing your child sick, and you don't know what's going on. You want them to take care of him. (ES 04)

Q20. **Original:** Ahí hablan español, pero a la hora que vas con el pediatra, él habla en inglés. Él a veces dice, si hay alguien para traducir, me va a traducir. Sino, él busca la manera de—cómo mi niño está grande, él entiende el inglés. Le dice, le chequea y todo, ya no trato, ya no hablo tanto y me pone el teléfono: "Todo está bien de tu niño y está todo bien," dice, "no hay ni una preocupación. Su peso está bien, sus medidas están bien, su vista está bien." Me pone en el teléfono y más al niño le habla, pues porque le hacen esto, le toman de los ojos, de leer lejos, que todo, que se deje el niño. De ahí estoy yo con él. Pero ya solo me dice, "todo está bien con el crecimiento." Me lo ponen en traductor. "No tengo intérprete y lo siento," dice. (ES04)
**Translation:** They speak Spanish there, but when you go to the pediatrician, he speaks in English. He says, if there's someone to translate, he'll translate for me. Otherwise, since my son is bigger now, he understands English. [The doctor] talks to him, he gives him his check-up and everything. I don't try anymore, I don't talk as much anymore, and [the doctor] uses the phone: "Everything is fine with your child and he's fine," he says, "There isn't a single concern. His weight is fine, his measurements are fine, his vision is fine." He uses the phone with me and talks more to my child, because they do this and that to him, they check his eyes, to read far away, everything [...]. I'm there with him. But now he just tells me, "Everything is fine with his growth." (ES04)

Q21. **Original:** A veces había cosas que yo decía, "tengo esta preocupación," pero pues ya mejor no la decía, porque se hacía como que un revuelo. Y en alguna ocasión me tocó una intérprete que también, como que, me regañaba. Me decía, y justo de ese tema de las de las botellas de de los biberones, me decía, "A ver, escucha, ya te dijo que no le des esa botella."
Y yo le decía, "es que no es la botella."
"No estás escuchando, que ya te dijo que no le des otra."
Entonces ni siquiera interpretaba. Ni siquiera interpretaba o pasaba el mensaje. Más bien, ella deducía lo que ya me había dicho, ella, y entonces me daba la indicación. Y entonces yo, de repente le decía, pero pasa el mensaje que yo le estoy diciendo. Yo empecé a tomar clases de inglés y después me di cuenta que ya no me estaba diciendo lo que la persona me quería comunicar, al final de esas sesiones que fueron, no sé, tal vez cinco. Ella dijo que yo no tomaba leche, que yo no consumía lácteos. Pero yo nunca dije eso. Y entonces la persona de aquí, la del la terapia, siempre creyó cosas que yo no hacía y que yo tampoco había dicho. Entonces, era como frustrante. Y era un caos esa situación. (ES21)
**Translation:** Sometimes there were things I would say to her, "I have this concern," but then I thought it would be better not to say it, because it would become a big deal. And on one occasion, I had an interpreter who also sort of scolded me. She told me, and precisely about that issue of baby bottle things, she would say, "Look, listen, she [medical provider] already told you not to give him that bottle."
And I would tell her, "It's not the bottle."
"You're not listening, she [medical provider] already told you not to give him another one."
So she didn't even interpret. She didn't even interpret or pass on the message. Rather, she would deduce what she [medical provider] had already told me, and then give me the instruction. And then I would suddenly say to her, "But pass on the message that I'm telling you."
I started taking English classes and then I realized that she was no longer telling me what the person wanted to communicate to me, at the end of those sessions, which were, I don't know, maybe five. She said that I didn't drink milk, that I didn't consume dairy products. But I never said that. And then the person here, the one in therapy, always believed things that I wasn't doing and that I hadn't said. So it was frustrating. And that situation was chaotic. (ES21)

Q22. **Original:** Confiar completamente, el número 5, porque realmente, si es algo grave. Tengo tres niños, y [si] hay algo grave, y está el intérprete allí, tengo que decirle lo que pasa del niño. No le puedo no contarle porque cómo lo va a traducir ella con el doctor. (ES04)
**Translation:** Trust completely, number 5, because really, I have two children, and if something serious is happening, and the interpreter is there, I have to tell her what's happening with the child. I can't not tell her, because how is she going to translate it for the doctor? (ES04)

Q23. **Original:** No apagaron la cámara y creyeron que era normal. [...] El personal se dio la vuelta y me dejaron a mí sola con el intérprete. [...] [El médico] no me estaba viendo a mí, sólo estaba viendo hacia el techo. (ES03)

**Translation:** They didn't turn off the camera, and they thought that was normal. [...] The medical staff turned around and left me alone with the interpreter. [...] The doctor was not looking at me. He was looking at the ceiling.

Q24. **Original:** Tenía personas ahí, se escuchaban, y que estaba peleando. Entonces, había más personas escuchando lo que yo decía. Entonces mi información, o sea, lo que yo le decía a la doctora, no era, No, no estaba siendo resguardada. [...] Ella me pidió disculpas cuando la doctora salió. Entonces, no debió haber pasado. De ahí perdí la confianza con la traducción. (ES03)
**Translation:** There were people there. They could hear each other, and she was fighting [with them]. So there were more people listening to what I was saying. So my information, that is, what I was saying to the doctor, wasn't being protected. [...] She apologized to me when the doctor left. So it shouldn't have happened. From then on, I lost confidence/trust in translation [interpretation]. (ES03)

Q25. **Original:** Todo el mundo le dice a uno que sí, que aquí guardamos bien tus datos, … pero nos estamos dando cuenta que uno hace una búsqueda en Google o uno escribe alguna cosa y a los 2 s ya te está apareciendo publicidad de lo que estabas buscando. Nos están escuchando en todos lados. Ya la tecnología nos tiene, mejor dicho, abordado. Entonces, yo pienso que la privacidad de la información, porque lo más valioso en este mundo es la información, el que tiene la información tiene el poder. Pero yo pienso que ya la verdad, mucha gente tiene acceso a toda nuestra información. No falte, sino que tengan acceso a nuestras cuentas bancarias, y yo creo que también ya la tienen. Así que ya eso no me preocupa. La verdad. (ES14)
**Translation:** Everyone tells you, yes, we keep your data safe here, … but we're noticing that if you say something or you do a Google search, or you type something, and within two seconds, you're already seeing ads for what you were looking for. They're listening to us everywhere. Technology has, better said, accosted us. So, I think that the privacy of information, because the most valuable thing in this world is information, whoever has the information has the power. But I think that, honestly, many people have access to all our information. They're only missing access to our bank accounts, and I think they already have that too. So, that doesn't worry me anymore. Honestly. (ES14)

Q26. **Original:** La verdad, No. Pienso que ya hoy la información es tan a todos que, no, ya todo mundo la tiene. (ES14)
**Translation:** To be honest, no. I think information is so widely available these days that, no, everyone has it now.

# Bibliography


[1] California Adopts New Employment AI Regulations Effective October 1, 2025 | Insights | Mayer Brown — mayerbrown.com. https://www.mayerbrown.com/en/insights/publications/2025/08/california-adopts-new-employment-ai-regulations-effective-october-1-2025. [Accessed 09-09-2025]. 3.7.4, 6.2.3

[2] State of Illinois 103rd General Assembly. Hb3773, 2023. URL https://www.ilga.gov/legislation/103/HB/10300HB3773.htm. 3.7.4, 6.2.3

[3] Basil Abraham, Danish Goel, Divya Siddarth, Kalika Bali, Manu Chopra, Monojit Choudhury, Pratik Joshi, Preethi Jyoti, Sunayana Sitaram, and Vivek Seshadri. Crowdsourcing speech data for low-resource languages from low-income workers. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Twelfth Language Resources and Evaluation Conference*, page 2819–2826, Marseille, France, May 2020. European Language Resources Association. ISBN 9791095546344. URL https://aclanthology.org/2020.lrec-1.343. 3.7.2

[4] Sara L. Ackerman, U. Sarkar, Lina Tieu, M. Handley, D. Schillinger, Kenneth J. Hahn, Mekhala Hoskote, Gato Gourley, and C. Lyles. Meaningful use in the safety net: a rapid ethnography of patient portal implementation at five community health centers in california. *J. Am. Medical Informatics Assoc.*, 2017. doi: 10.1093/jamia/ocx015. 5.3.3

[5] Devon Adams, Alseny Bah, Catherine Barwulor, Nureli Musaby, Kadeem Pitkin, and Elissa M. Redmiles. Ethics emerging: the story of privacy and security perceptions in virtual reality. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, 2018. ISBN 978-1-939133-10-6. URL https://www.usenix.org/conference/soups2018/presentation/adams. 2.4.1 (b)

[6] Shimaa Ahmed, Amrita Roy Chowdhury, Kassem Fawaz, and Parmesh Ramanathan. Preech: A system for Privacy-Preserving speech transcription. page 2703–2720, 2020. ISBN 9781939133175. URL https://www.usenix.org/conference/usenixsecurity20/presentation/ahmed-shimaa. 3.7.4

[7] Alëna Aksënova, Daan van Esch, James Flynn, and Pavel Golik. How might we create better benchmarks for speech recognition? In Kenneth Church, Mark Liberman, and Valia Kordoni, editors, *Proceedings of the 1st Workshop on Benchmarking: Past, Present and Future*, pages 22–34, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.bppf-1.4. URL https://aclanthology.org/2021.bppf-1.4/. 3.7.2

[8] Alëna Aksënova, Antoine Bruguier, Amanda Ritchart-Scott, and Uri Mendlovic. Algorithmic exploration of american english dialects. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7374–7378, 2020. doi: 10.1109/ICASSP40776.2020.9053751. 3.4.4

[9] Alëna Aksënova, Zhehuai Chen, Chung-Cheng Chiu, Daan van Esch, Pavel Golik, Wei Han, Levi King, Bhuvana Ramabhadran, Andrew Rosenberg, Suzan Schwartz, and Gary Wang. Accented speech recognition: Benchmarking, pre-training, and diverse data, 2022. URL https://arxiv.org/abs/2205.08014. 3.7.2

[10] Bushra A. Alahmadi, Louise Axon, and Ivan Martinovic. 99% False Positives: A Qualitative Study of SOC Analysts' Perspectives on Security Alarms. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 2783–2800. USENIX Association, August 2022. ISBN 978-1-939133-31-1. URL https://www.us

enix.org/conference/usenixsecurity22/presentation/alahmadi. 4.4.3

[11] Mohammed Alghassab. Analyzing the Impact of Cybersecurity on Monitoring and Control Systems in the Energy Sector. *Energies*, 15(1):218, January 2022. doi: 10.3390/en15010218. URL https://www.mdpi.com/1996-1073/15/1/218. 4.4.2

[12] Sami Alkhatib, Ryan Kelly, Jenny Waycott, George Buchanan, Marthie Grobler, and Shuo Wang. "who wants to know all this stuff?!": Understanding older adults' privacy concerns in aged care monitoring devices. *Interacting with Computers*, 33(5):481–498, 2021. doi: 10.1093/itnow/bwac034. 2.4.4

[13] Rachael Meghan Allbritten. *Sounding Southern: phonetic features and dialect perceptions*. thesis, Georgetown University, 2011. URL https://repository.library.georgetown.edu/handle/10822/553139. 3.4.4

[14] Anita L. Allen. *Unpopular Privacy: What Must We Hide?* Studies in feminist philosophy. Oxford University Press, New York, 2011. ISBN 1-283-42379-0. 2.4.4

[15] Nazanin Andalibi and Justin Buss. The human in emotion recognition on social media: Attitudes, outcomes, risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–16, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. doi: 10.1145/3313831.3376680. URL https://doi.org/10.1145/3313831.3376680. 3.7.4

[16] Laurie M Anderson, Susan C Scrimshaw, Mindy T Fullilove, Jonathan E Fielding, and Jacques Normand. Culturally competent healthcare systems. *American Journal of Preventive Medicine*, 24(3):68–79, 2003. doi: 10.1016/S0749-3797(02)00657-8. URL https://linkinghub.elsevier.com/retrieve/pii/S0749379702006578. 5.2

[17] Claudia V. Angelelli. *Interpreters' voices*, page 105–128. Cambridge University Press, Cambridge, 2004. ISBN 9780521066778. doi: 10.1017/CBO9780511486616.008. URL https://www.cambridge.org/core/books/medical-interpreting-and-crosscultural-communication/interpreters-voices/44E66B354098A42EFBD8538184B4D386. 5.3.3

[18] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine bias, May 2016. URL https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. 3.7.3

[19] Marcy G. Antonio, O. Petrovskaya, and Francis Y. Lau. Is research on patient portals attuned to health equity? a scoping review. *Journal of the American Medical Informatics Association*, 2019. doi: 10.1093/jamia/ocz054. 5.3.3

[20] George E. Apostolakis and Douglas M. Lemon. A Screening Methodology for the Identification and Ranking of Infrastructure Vulnerabilities Due to Terrorism. *Risk Analysis*, 25(2):361–376, 2005. doi: 10.1111/j.1539-6924.2005.00595.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2005.00595.x. 4.4.2

[21] Apple. Legal - apple privacy policy - apple, 2022. URL https://www.apple.com/legal/privacy/en-ww/. Accessed June 9, 2023. 2.4.1 (b)

[22] Apple. Legal - data & privacy - apple, 2023. URL https://www.apple.com/legal/privacy/data/. Accessed June 9, 2023. 2.4.1 (b)

[23] Apple. Introducing apple vision pro, 2023. URL https://www.apple.com/newsroom/2023/06/introducing-apple-vision-pro/. 2.4.1 (a), 2.4.1 (b)

[24] Noah Apthorpe, Sarah Varghese, and Nick Feamster. Evaluating the contextual integrity of privacy regulation: Parents' IoT toy privacy norms versus COPPA. In *28th USENIX Security Symposium (USENIX Security 19)*, 2019. ISBN 978-1-939133-06-9. URL https://www.usenix.org/conference/usenixsecurity19/presentation/apthorpe. 2.4.4

[25] Lena Armstrong, Abbey Liu, Stephen MacNeil, and Danaë Metaxa. The silicon ceiling: Auditing gpt's race and gender biases in hiring. In *Proceedings of the 4th ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, pages 1–18, 2024. 3.4.1

[26] Miriam E. Armstrong, Keith S. Jones, Akbar Siami Namin, and David C. Newton. The Knowledge, Skills, and Abilities Used by Penetration Testers: Results of Interviews with Cybersecurity Professionals in Vulnerability Assessment and Management. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62(1):709–713, 2018. doi: 10.1177/1541931218621161. URL http://journals.sagepub.com/doi/10.1177/1541931218621161. 4.4.3

[27] Miriam E. Armstrong, Keith S. Jones, Akbar Siami Namin, and David C. Newton. Knowledge, Skills, and Abilities for Specialized Curricula in Cyber Defense: Results from Interviews with Cyber Professionals. *ACM Transactions on Computing Education*, 20(4):29:1–29:25, November 2020. doi: 10.1145/3421254. URL https://doi.org/10.1145/3421254. 4.4.3

[28] Arif Bacchus. Microsoft hololens 2 hands-on review. *Digital Trends*, Nov 2019. URL https://www.digitaltrends.com/computing/microsoft-hololens-2-ar-hands-on-features-price-photos-video-release-date/. Accessed November 30, 2022. 2.4.1 (a)

[29] Leen Bakdash, Areeba Abid, Amritha Gourisankar, and Tracey L. Henry. Chatting Beyond ChatGPT: Advancing Equity Through AI-Driven Language Interpretation. *Journal of General Internal Medicine*, 39(3): 492–495, February 2024. doi: 10.1007/s11606-023-08497-6. URL https://doi.org/10.1007/s11606-023-08497-6. 5.3.2

[30] Mary Elizabeth Ballard and Kelly Marie Welch. Virtual warfare: Cyberbullying and cyber-victimization in mmog play. *Games and Culture*, 12(5), 2017. doi: 10.1177/1555412015592473. 2.4.2 (b)

[31] Nathan Barrett, Andrew McEachin, Jonathan N Mills, and Jon Valant. Disparities and discrimination in student discipline by race and family income. *Journal of Human Resources*, 56(3):711–748, 2021. 3.4.1

[32] Muzakki Bashori, Roeland van Hout, Helmer Strik, and Catia Cucchiarini. 'look, i can speak correctly': learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*, 37(5-6):1335–1363, 2024. 3.3, 3.4.1, 3.4.3

[33] Jeanne Batalova and Michael Fix. *New Brain Gain: Rising Human Capital among Recent Immigrants to the United States*. URL https://www.migrationpolicy.org/research/new-brain-gain-rising-human-capital-among-recent-immigrants-united-states. 5.3.1

[34] Amanda Baughan, Xuezhi Wang, Ariel Liu, Allison Mercurio, Jilin Chen, and Xiao Ma. A mixed-methods approach to understanding user trust after voice assistant failures. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, page 1–16, Hamburg Germany, April 2023. ACM. ISBN 9781450394215. doi: 10.1145/3544548.3581152. URL https://dl.acm.org/doi/10.1145/3544548.3581152. 3.4.3, 5.3.6

[35] Scott Beach, Richard Schulz, Julie Downs, Judith Matthews, Bruce Barron, and Katherine Seelman. Disability, age, and informational privacy attitudes in quality of life technology applications: Results from a national web survey. *ACM Transactions on Accessible Computing*, 2(1):5:1–5:21, May 2009. doi: 10.1145/1525840.1525846. URL https://doi.org/10.1145/1525840.1525846. 3.5.4

[36] Kate Behncken. Closing the cybersecurity skills gap –microsoft expands efforts to 23 countries. https://blogs.microsoft.com/blog/2022/03/23/closing-the-cybersecurity-skills-gap-microsoft-expands-efforts-to-23-countries/, March 2022. 4.3

[37] Genevieve Bell, Mark Blythe, and Phoebe Sengers. Making by making strange: Defamiliarization and the design of domestic technologies. *ACM Trans. Comput.-Hum. Interact.*, 12(2):149–173, jun 2005. ISSN 1073-0516. doi: 10.1145/1067860.1067862. URL https://doi.org/10.1145/1067860.1067862. 3.7.2

[38] Luke Bencie and Sami Araboghli. A 6-Part Tool for Ranking and Assessing Risks. *Harvard Business Review*, September 2018. https://hbr.org/2018/09/a-6-part-tool-for-ranking-and-assessing-risks. 4.6.1 (a)

[39] Veroniek Binkhorst, Tobias Fiebig, Katharina Krombholz, Wolter Pieters, and Katsiaryna Labunets. Security at the End of the Tunnel: The Anatomy of VPN Mental Models Among Experts and Non-Experts in a Corporate Context. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 3433–3450. USENIX Association, August 2022. ISBN 9781939133311. URL https://www.usenix.org/conference/usen

ixsecurity22/presentation/binkhorst. 4.4.3

[40] Galina B. Bolden. Toward understanding practices of medical interpreting: Interpreters' involvement in history taking. *Discourse Studies*, 2(4):387–419, 2000. doi: 10.1177/1461445600002004001. URL https://journals.sagepub.com/doi/10.1177/1461445600002004001. 5.3.3, 5.3.8

[41] Seppo Borenius, Pavithra Gopalakrishnan, Lina Bertling Tjernberg, and Raimo Kantola. Expert-Guided Security Risk Assessment of Evolving Power Grids. *Energies*, 15(9):3237, January 2022. doi: 10.3390/en15093237. URL https://www.mdpi.com/1996-1073/15/9/3237. 4.4.2

[42] David Botta, Rodrigo Werlinger, André Gagné, Konstantin Beznosov, Lee Iverson, Sidney Fels, and Brian Fisher. Towards Understanding IT Security Professionals and Their Tools. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, SOUPS '07, pages 100–111. Association for Computing Machinery, 2007. ISBN 9781595938015. doi: 10.1145/1280680.1280693. URL https://doi.org/10.1145/1280680.1280693. 4.4.3

[43] Russell Brandom, Viola Zhou, and Sanghamitra Kar P. The AI job interviewer will see you now. *Rest of World*, July 2024. URL https://restofworld.org/2024/ai-interview-software-hiring-practices/. 3.3

[44] Dale E. Brashers, Daena J. Goldsmith, and Elaine Hsieh. Information seeking and avoiding in health contexts. *Human Communication Research*, 28(2):258–271, 2002. doi: 10.1111/j.1468-2958.2002.tb00807.x. URL https://academic.oup.com/hcr/article/28/2/258-271/4331134. 5.3.8

[45] Robin N. Brewer, Christina Harrington, and Courtney Heldreth. Envisioning equitable speech technologies for black older adults. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '23, page 379–388, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400701924. doi: 10.1145/3593013.3594005. URL https://doi.org/10.1145/3593013.3594005. 3.4.3

[46] Kristin L. Bryan, Christina Lamoureux, and Dan Lonergan. 2021 Year in Review: Biometric and AI Litigation. *The National Law Review*, January 2022. URL https://www.natlawreview.com/article/2021-year-review-biometric-and-ai-litigation. 2.9.5

[47] Kent Bye, Diane Hosfelt, Sam Chase, Matt Miesnieks, and Taylor Beck. The ethical and privacy implications of mixed reality. In *ACM SIGGRAPH 2019 Panels*, 2019. doi: 10.1145/3306212.3328138. 2.4.2

[48] Eric Byres. The air gap: SCADA's enduring security myth. *Commun. ACM*, 56(8):29–31, August 2013. ISSN 0001-0782. doi: 10.1145/2492007.2492018. URL https://doi.org/10.1145/2492007.2492018. 4.3

[49] Veena Calambur, Dongwhan Jun, Melody K Schiaffino, Zhan Zhang, and Jina Huh-Yoo. A case for "little english" in nurse notes from the telehealth intervention program for seniors: Implications for future design and research. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, page 1–16, Honolulu HI USA, May 2024. ACM. ISBN 9798400703300. doi: 10.1145/3613904.3641961. URL https://dl.acm.org/doi/10.1145/3613904.3641961. 5.2

[50] Jim Carlton and Paul Overberg. Immigration is the only thing propping up california's population. *The Wall Street Journal*, May 2025. URL https://www.wsj.com/business/california-population-growth-immigration-h-1b-visa-4b526478. 5.3.1

[51] Olveen Carrasquillo, E. John Orav, Troyen A. Brennan, and Helen R. Burstin. Impact of language barriers on patient satisfaction in an emergency department. *Journal of General Internal Medicine*, 14(2):82–87, 1999. doi: 10.1046/j.1525-1497.1999.00293.x. URL http://link.springer.com/10.1046/j.1525-1497.1999.00293.x. 5.3.4

[52] Alejandra Casillas, Giselle Perez-Aguilar, Anshu Abhat, Griselda Gutierrez, Tanya T. Olmos-Ochoa, Carmen Mendez, Anish P. Mahajan, Arleen F. Brown, and Gerardo Moreno. Su salud a la mano (your health at hand): patient perceptions about a bilingual patient portal in the los angeles safety net. *Journal of the American Medical Informatics Association*, 2019. doi: 10.1093/jamia/ocz115. 5.3.3

[53] CDC. Health Insurance Portability and Accountability Act of 1996 (HIPAA), September 2024. URL https://www.cdc.gov/phlp/php/resources/health-insurance-portability-and-accountability-a

ct-of-1996-hipaa.html. 5.3.2

[54] The Immigrant Learning Center. The impact of immigrants on health care in the united states. URL https://www.ilctr.org/about-immigrants/ilc-publications-and-resources/the-impact-of-immigrants-on-health-care-in-the-united-states/. 5.3.1

[55] Ella Chakarian. Welcome to the job search. your AI recruiter will see you now. *The San Francisco Standard*, March 2025. URL https://sfstandard.com/2025/03/31/zara-ai-recruiter-job-interviews/. 3.3

[56] Kalvin Chang, Yi-Hui Chou, Jiatong Shi, Hsuan-Ming Chen, Nicole Holliday, Odette Scharenborg, and David R. Mortensen. Self-supervised Speech Representations Still Struggle with African American Vernacular English, August 2024. URL http://arxiv.org/abs/2408.14262. arXiv:2408.14262 [cs, eess]. 3.4.3, 5.3.6

[57] Te-Ping Chen, Kevin Hand, and Yan Wu. Tech that aims to improve meetings. *Wall Street Journal*, Jan 2021. URL https://www.wsj.com/articles/tech-that-aims-to-improve-meetings-11610640133. Accessed November 30, 2022. 2.4.1 (a)

[58] Xuewei Chen, E. Schofield, J. Hay, Erika A. Waters, M. Kiviniemi, and H. Orom. Race/ethnicity, nativity status, and patient portal access and use. *Journal of health care for the poor and underserved*, 2022. doi: 10.1353/hpu.2021.0099. 5.3.3

[59] Zhisheng Chen. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, 10(1):567, September 2023. doi: 10.1057/s41599-023-02079-x. URL https://www.nature.com/articles/s41599-023-02079-x. 3.3

[60] Yulia Cherdantseva, Pete Burnap, Andrew Blyth, Peter Eden, Kevin Jones, Hugh Soulsby, and Kristan Stoddart. A review of cyber security risk assessment methods for SCADA systems. *Computers & Security*, 56: 1–27, February 2016. doi: 10.1016/j.cose.2015.09.009. URL https://www.sciencedirect.com/science/article/pii/S0167404815001388. 4.4.2

[61] Chola Chhetri and Vivian Motti. "I mute my echo when I talk politics": Connecting Smart Home Device Users' Concerns to Privacy Harms Taxonomy. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 66. SAGE Publications, 2022. doi: 10.1177/1071181322661114. 2.4.4

[62] Becky Childs and Christine Mallinson. *Southern Language and Linguistics*, page 54–65. Routledge, 1 edition, 2017. URL https://www.taylorfrancis.com/chapters/edit/10.4324/9781315768076-5/southern-language-linguistics-becky-childs-christine-mallinson. 3.4.4

[63] Jinhyun Cho. Interpreters as translation machines: Telephone interpreting challenges as awareness problems. *Qualitative Health Research*, 33(12):1037–1048, 2023. doi: 10.1177/10497323231191712. URL https://journals.sagepub.com/doi/10.1177/10497323231191712. 5.3.3

[64] Janet N. Chu, Urmimala Sarkar, Natalie A. Rivadeneira, Robert A. Hiatt, and Elaine C. Khoong. Impact of language preference and health literacy on health information-seeking experiences among a low-income, multilingual cohort. *Patient Education and Counseling*, 105(5):1268–1275, 2022. ISSN 0738-3991. doi: https://doi.org/10.1016/j.pec.2021.08.028. URL https://www.sciencedirect.com/science/article/pii/S0738399121005747. 5.3.4

[65] Phoebe K Chua, Hillary Abraham, and Melissa Mazmanian. Playing the hiring game: Class-based emotional experiences and tactics in elite hiring. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2): 1–27, 2021. 3.4.1

[66] Ching-Hua Chuan, Wan-Hsiu S. Tsai, Di Lun, and Nicholas Carcioppolo. *Understanding Responses to Embarrassing Questions in Chatbot-Facilitated Medical Interview Conversations Using Deep Language Models*, page 17–25. Springer Nature Switzerland, Cham, 2024. ISBN 9783031344589. doi: 10.1007/978-3-031-34459-6_2. URL https://link.springer.com/10.1007/978-3-031-34459-6_2. 5.3.8

[67] CISA. US-CERT and ICS-CERT Transition to CISA | CISA. https://www.cisa.gov/news-events/alerts/2023/02/24/us-cert-and-ics-cert-transition-cisa, February 2023. 4.6.1 (a)

[68] CISA. Cybersecurity Alerts & Advisories | CISA. https://www.cisa.gov/news-events/cybersecurity-advisories, 2024. Accessed February 22, 2024. 4.6.1 (a)

[69] Michelle Clark and Sharon Bailey. *Chatbots in Health Care: Connecting Patients to Information: Emerging Health Technologies*. CADTH Horizon Scans. Canadian Agency for Drugs and Technologies in Health, Ottawa (ON), 2024. URL http://www.ncbi.nlm.nih.gov/books/NBK602381/. 5.3.8

[70] Cole Claybourn. Is AI affecting college admissions? *U.S. News & World Report*, December 2023. URL https://www.usnews.com/education/best-colleges/articles/is-ai-affecting-college-admissions. 3.3

[71] European Commission. Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, 2021. URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206. COM/2021/206 final. 3.4.1

[72] Office of the Commissioner. 21st century cures act, January 2020. URL https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act. 5.3.2

[73] Wm. Arthur Conklin. IT vs. OT Security: A Time to Consider a Change in CIA to Include Resilience. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 2642–2647, January 2016. doi: 10.1109/HICSS.2016.331. URL https://doi.org/10.1109/HICSS.2016.331. 4.4.1

[74] Jo Constantz. Job interviews enter a strange new world with AI that talks back. *Bloomberg*, May 2025. URL https://www.bloomberg.com/news/articles/2025-05-28/job-applicant-interviews-conducted-by-ai-offer-benefits-tech-glitches. 3.3

[75] HTC Corporation. Vive Legal Documents | HTC United States, 2022. URL https://www.htc.com/us/terms/vive/. Accessed November 30, 2022. 2.4.1 (b)

[76] HTC Corporation. Privacy Policy | Terms | HTC United States, 2022. URL https://www.htc.com/us/terms/privacy/. Accessed November 30, 2022. 2.4.1 (b)

[77] The MITRE Corporation. Home | CVE. https://www.cve.org/, 2024. Accessed February 22, 2024. 4.6.1 (a)

[78] The MITRE Corporation. CWE - Common Weakness Enumeration. https://cwe.mitre.org/, 2024. Accessed February 22, 2024. 4.6.1 (a)

[79] The MITRE Corporation. MITRE ATT&CK. https://attack.mitre.org/, 2024. Accessed February 22, 2024. 4.6.1 (a)

[80] Amanda C. Cote. "I Can Defend Myself": Women's Strategies for Coping With Harassment While Gaming Online. *Games and Culture*, 12(2), 2017. doi: 10.1177/1555412015587603. 2.4.2 (b)

[81] American Immigration Council. Immigrants in the united states, 2025. URL https://map.americanimmigrationcouncil.org/locations/national/#. 5.3.1

[82] Council of European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). *Official Journal of the European Union*, (2024/1689), June 2024. URL http://data.europa.eu/eli/reg/2024/1689/oj/eng. 3.7.4, 6.2.3

[83] Jennifer Suzanne Cramer. *The effect of borders on the linguistic production and perception of regional identity in Louisville, Kentucky*. University of Illinois at Urbana-Champaign, 2010. 3.4.4

[84] Julie Creswell, Nicole Perlroth, and Noam Scheiber. Ransomware Disrupts Meat Plants in Latest Attack on Critical U.S. Business. *The New York Times*, June 2021. https://www.nytimes.com/2021/06/01/business/meat-plant-cyberattack-jbs.html. 4.3

[85] Roderic Crooks and Morgan Currie. Numbers will not save us: Agonistic data practices. *The Information Society*, 37(4), 2021. doi: 10.1080/01972243.2021.1920081. URL https://par.nsf.gov/biblio/102

84249. 3.4.2

[86] Alvaro A. Cárdenas, Saurabh Amin, Zong-Syun Lin, Yu-Lun Huang, Chi-Yen Huang, and Shankar Sastry. Attacks against process control systems: risk assessment, detection, and response. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '11, pages 355–366. Association for Computing Machinery, March 2011. ISBN 9781450305648. doi: 10.1145/1966913.1966959. URL https://doi.org/10.1145/1966913.1966959. 4.4.2

[87] Jonathan Dalby and Diane Kewley-Port. Explicit pronunciation training using automatic speech recognition technology. *CALICO journal*, pages 425–445, 1999. 3.3, 3.4.1, 3.4.3

[88] Jeffrey Dastin. Insight - amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, October 2018. URL https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/. 3.3

[89] K Davis, R Berthier, S Zonouz, G Weaver, R Bobba, E Rogers, P Sauer, and D Nicol. Cyber-physical security assessment (CYPSA) for electric power systems. *IEEE-HKN: THE BRIDGE*, 2016. 4.4.2

[90] Seethalakshmi H. Davis, Julia Rosenberg, Jenny Nguyen, Manuel Jimenez, K. Casey Lion, Gabriela Jenicek, Harry Dallmann, and Katherine Yun. Translating discharge instructions for limited english-proficient families: Strategies and barriers. *Hospital Pediatrics*, 9(10):779–787, October 2019. doi: 10.1542/hpeds.2019-0055. 5.3.3

[91] Jaybie A. De Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. Security and privacy approaches in mixed reality: A literature survey. *ACM Comput. Surv.*, 52(6), 2019. doi: 10.1145/3359626. 2.4.2

[92] Daniel Delmonaco, Samuel Mayworm, Hibby Thach, Josh Guberman, Aurelia Augusta, and Oliver L. Haimson. "What are you doing, TikTok?": How Marginalized Social Media Users Perceive, Theorize, and "Prove" Shadowbanning. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW1), April 2024. doi: 10.1145/3637431. URL https://doi.org/10.1145/3637431. 3.4.2

[93] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. In situ with bystanders of augmented reality glasses: Perspectives on recording and privacy-mediating technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, 2014. doi: 10.1145/2556288.2557352. 2.4.2, 2.4.2 (a), 2.9.4

[94] Patrick Denzler and Wolfgang Kastner. *Reference Architectures for Closing the IT/OT Gap*, pages 95–123. Springer, 2023. ISBN 9783662650035 9783662650042. doi: 10.1007/978-3-662-65004-2_4. URL https://link.springer.com/10.1007/978-3-662-65004-2_4. 4.4.1

[95] Jayati Dev and Sruti Dev. "how can i help you?": User perceptions of privacy in retail chat agents. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, page 1–6, Hamburg Germany, April 2023. ACM. ISBN 9781450394222. doi: 10.1145/3544549.3585796. URL https://dl.acm.org/doi/10.1145/3544549.3585796. 5.3.8

[96] Michael Ann DeVito. How transfeminine tiktok creators navigate the algorithmic trap of visibility via folk theorization. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2), November 2022. doi: 10.1145/3555105. URL https://doi.org/10.1145/3555105. 3.4.2

[97] Alicia DeVos, Aditi Dhabalia, Hong Shen, Kenneth Holstein, and Motahhare Eslami. Toward User-Driven Algorithm Auditing: Investigating users' strategies for uncovering harmful algorithmic behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. doi: 10.1145/3491102.3517441. URL https://doi.org/10.1145/3491102.3517441. 3.7.1, 3.7.2

[98] Kristin N Dew, Anne M Turner, Yong K Choi, Alyssa Bosold, and Katrin Kirchhoff. Development of machine translation technology for assisting health communication: A systematic review. *Journal of biomedical informatics*, 85:56–67, 2018. 5.3.5

[99] Payal Dhar. Cybersecurity Report: "Smart Farms" Are Hackable Farms. *IEEE Spectrum*, March 2021. https://spectrum.ieee.org/cybersecurity-report-how-smart-farming-can-be-hacked. 4.3

[100] Lisa Diamond, Karen Izquierdo, Dana Canfield, Konstantina Matsoukas, and Francesca Gany. A systematic review of the impact of patient–physician non-english language concordance on quality of care and outcomes. *Journal of General Internal Medicine*, 34(8):1591–1606, 2019. doi: 10.1007/s11606-019-04847-5. URL http://link.springer.com/10.1007/s11606-019-04847-5. 5.3.4

[101] Lisa C. Diamond, Amy Wilson-Stronks, and Elizabeth A. Jacobs. Do hospitals measure up to the national culturally and linguistically appropriate services standards? *Medical Care*, 48(12):1080–1087, 2010. doi: 10.1097/MLR.0b013e3181f380bc. URL https://journals.lww.com/00005650-201012000-00006. 5.3.3

[102] Bryan Dosono and Bryan Semaan. Decolonizing tactics as collective resilience: Identity work of aapi communities on reddit. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW1), May 2020. doi: 10.1145/3392881. URL https://doi.org/10.1145/3392881. 3.7.1

[103] Jonathan Downie and Angela Dickson. Unsound evaluations of medical machine translation risk patient health and confidentiality. *JAMA Internal Medicine*, 179(7):1001–1001, 07 2019. ISSN 2168-6106. doi: 10.1001/jamainternmed.2019.1856. URL https://doi.org/10.1001/jamainternmed.2019.1856. 5.3.2, 5.3.5

[104] John J. Dudley, Jason T. Jacques, and Per Ola Kristensson. Crowdsourcing design guidance for contextual adaptation of text content in augmented reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, May 2021. ISBN 978-1-4503-8096-6. doi: 10.1145/3411764.3445493. URL https://doi.org/10.1145/3411764.3445493. 2.9.5

[105] Sarah Dunn. What brings you here today?: Challenges for clinicians communicating data to patients. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, page 1–7, Yokohama Japan, April 2025. ACM. ISBN 9798400713958. doi: 10.1145/3706599.3719283. URL https://dl.acm.org/doi/10.1145/3706599.3719283. 5.2

[106] Chandelis Duster. Energy secretary says adversaries have capability of shutting down US power grid | CNN Politics. *CNN*, June 2021. https://www.cnn.com/2021/06/06/politics/us-power-grid-jennifer-granholm-cnntv/index.html. 4.3

[107] New American Economy. *Taxes & Spending Power: How immigration plays a critical role*. New American Economy, 2025. URL https://www.newamericaneconomy.org/issues/taxes-spending-power/. 5.3.1

[108] Lisa Egede, Leslie Coney, Brittany Johnson, Christina Harrington, and Denae Ford. "For Us By Us": Intentionally Designing Technology for Lived Black Experiences. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, DIS '24, page 3210–3224, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400705830. doi: 10.1145/3643834.3661535. URL https://doi.org/10.1145/3643834.3661535. 3.7.2

[109] Ben Egliston and Marcus Carter. "the metaverse and how we'll build it": The political economy of Meta's Reality Labs. *New Media & Society*, 2022. doi: 10.1177/14614448221119785. 2.9.5

[110] Jacob Eisenstein, Vinodkumar Prabhakaran, Clara Rivera, Dorottya Demszky, and Devyani Sharma. Md3: The multi-dialect dataset of dialogues, 2023. URL https://arxiv.org/abs/2305.11355. 3.4.3

[111] Amal Khalil Elimat and Ali Farhan AbuSeileek. Automatic speech recognition technology as an effective means for teaching pronunciation. *Jalt Call Journal*, 10(1):21–47, 2014. 3.3, 3.4.1, 3.4.3

[112] Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. Exploring how privacy and security factor into iot device purchase behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019. doi: 10.1145/3290605.3300764. 2.6

[113] Pardis Emami-Naeini, Yuvraj Agarwal, Lorrie Faith Cranor, and Hanan Hibshi. Ask the Experts: What Should Be on an IoT Privacy and Security Label? In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 447–464, May 2020. doi: 10.1109/SP40000.2020.00043. 4.4.3

[114] Robert Espinoza. Building a stronger workforce means embracing immigrant contributions, February 2025. URL https://nationalskillscoalition.org/blog/news/building-a-stronger-workforce-m

eans-embracing-immigrant-contributions/. 5.3.1

[115] Fernando M. Espinoza-Cuadros, Juan M. Perero-Codosero, Javier Antón-Martín, and Luis A. Hernández-Gómez. Speaker de-identification system using autoencoders and adversarial training, Nov 2020. URL http://arxiv.org/abs/2011.04696. 3.7.4

[116] Rose Eveleth. Google glass wasn't a failure. it raised crucial concerns. *Wired*, 2018. ISSN 1059-1028. URL https://www.wired.com/story/google-glass-reasonable-expectation-of-privacy/. 2.3

[117] Alessandro Fabris, Nina Baranowska, Matthew J. Dennis, David Graus, Philipp Hacker, Jorge Saldivar, Frederik Zuiderveen Borgesius, and Asia J. Biega. Fairness and bias in algorithmic hiring: A multidisciplinary survey. *ACM Trans. Intell. Syst. Technol.*, 16(1), January 2025. ISSN 2157-6904. doi: 10.1145/3696457. URL https://doi.org/10.1145/3696457. 3.3, 3.4.1

[118] Cori Faklaris, Asa Blevins, Matthew O'Haver, Neha Singhal, and Francesco Cafaro. An exploration of user and bystander attitudes about mobile live-streaming video. 2017. doi: 10.13140/RG.2.2.14052.22406. URL http://arxiv.org/abs/1902.06671. 2.4.2 (c)

[119] Ben Falchuk, Shoshana Loeb, and Ralph Neff. The social metaverse: Battle for privacy. *IEEE Technology and Society Magazine*, 37(2), 2018. doi: 10.1109/MTS.2018.2826060. 2.4.2 (b)

[120] Spencer Fane. AI in hiring: The impact of illinois' new law on employment decisions and what businesses need to know, September 2024. URL https://www.mondaq.com/unitedstates/discrimination-disability-sexual-harassment/1519132/ai-in-hiring-the-impact-of-illinois-new-law-on-employment-decisions-and-what-businesses-need-to-know. 3.7.4, 6.2.3

[121] Siyuan Feng, Bence Mark Halpern, Olya Kudina, and Odette Scharenborg. Towards inclusive automatic speech recognition. *Computer Speech & Language*, 84:101567, March 2024. doi: 10.1016/j.csl.2023.101567. URL https://www.sciencedirect.com/science/article/pii/S0885230823000864. 3.4.3, 5.3.6

[122] Guo Freeman, Divine Maloney, Dane Acena, and Catherine Barwulor. (re)discovering the physical body online: Strategies and challenges to approach non-cisgender identity in social virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, Apr 2022. doi: 10.1145/3491102.3502082. URL https://doi.org/10.1145/3491102.3502082. 2.4.2 (b)

[123] Batya Friedman and Helen Nissenbaum. Bias in computer systems. *ACM Trans. Inf. Syst.*, 14(3):330–347, July 1996. ISSN 1046-8188. doi: 10.1145/230538.230561. URL https://doi.org/10.1145/230538.230561. 3.4.2

[124] Yang Gao, Yincheng Jin, Jagmohan Chauhan, Seokmin Choi, Jiyang Li, and Zhanpeng Jin. Voice in ear: Spoofing-resistant and passphrase-independent body sound authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(1), 2021. doi: 10.1145/3448113. 2.4.1 (a)

[125] Daniela Garcia-Castillo and Michael Fetters. Quality in medical translations: A review. *Journal of Health Care for the Poor and Underserved*, 18(1):74–84, 2007. ISSN 1548-6869. URL https://muse.jhu.edu/pub/1/article/210741. Johns Hopkins University Press. 5.3.5

[126] Shefali Garg, Zhouyuan Huo, Khe Chai Sim, Suzan Schwartz, Mason Chua, Alëna Aksënova, Tsendsuren Munkhdalai, Levi King, Darryl Wright, Zion Mengesha, Dongseong Hwang, Tara Sainath, Françoise Beaufays, and Pedro Moreno Mengibar. Improving Speech Recognition for African American English with Audio Classification. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 12356–12360, 2024. doi: 10.1109/ICASSP48485.2024.10447116. 3.7.2

[127] Tanmay Garg, Sarah Masud, Tharun Suresh, and Tanmoy Chakraborty. Handling bias in toxic speech detection: A survey. *ACM Computing Surveys*, 55(13s):264:1–264:32, July 2023. doi: 10.1145/3580494. URL https://doi.org/10.1145/3580494. 3.4.4

[128] Kathrin Gerling and Katta Spiel. A critical examination of virtual reality technology in the context of the minority body. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. doi: 10.1145/3411764.3445196. 2.4.2 (b)

[129] Alexandra Reeve Givens, Hilke Schellmann, and Julia Stoyanovich. We need laws to take on racism and

sexism in hiring technology. *The New York Times*, March 2021. URL https://www.nytimes.com/2021/03/17/opinion/ai-employment-bias-nyc.html. 3.3

[130] Cornelius Glackin, Gerard Chollet, Nazim Dugan, Nigel Cannings, Julie Wall, Shahzaib Tahir, Indranil Ghosh Ray, and Muttukrishnan Rajarajan. Privacy preserving encrypted phonetic search of speech data. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, page 6414–6418, Mar 2017. doi: 10.1109/ICASSP.2017.7953391. 3.7.4

[131] N. S. Goedhart, T. Zuiderent-Jerak, Joey Woudstra, J. Broerse, A. W. Betten, and C. Dedding. Persistent inequitable design and implementation of patient portals for users at the margins. *J. Am. Medical Informatics Assoc.*, 2021. doi: 10.1093/jamia/ocaa273. 5.3.3

[132] Simon B. Goldberg, Michael Tanana, Zac E. Imel, David C. Atkins, Clara E. Hill, and Timothy Anderson. Can a computer detect interpersonal skills? using machine learning to scale up the facilitative interpersonal skills task. *Psychotherapy Research*, 31(3):281–288, April 2021. doi: 10.1080/10503307.2020.1741047. URL https://www.tandfonline.com/doi/full/10.1080/10503307.2020.1741047. 3.3, 3.4.1

[133] Google. Glass terms of use, 2014. URL https://www.google.com/glass/termsofuse/. Accessed November 30, 2022. 2.4.1 (b)

[134] Google. Wearable emotion detection and feedback system - google patents, Mar 2015. URL https://patents.google.com/patent/US9508008B2/en. 2.4.1 (a)

[135] Google. Google privacy policy, Oct 2022. URL https://policies.google.com/privacy. Accessed November 30, 2022. 2.4.1 (b)

[136] Stephen D.N. Graham. Software-sorted geographies. *Progress in Human Geography*, 29(5):562–580, 2005. doi: 10.1191/0309132505ph568oa. URL https://doi.org/10.1191/0309132505ph568oa. 3.4.2

[137] Matthew Groh, Caleb Harris, Roxana Daneshjou, Omar Badri, and Arash Koochek. Towards transparency in dermatology image datasets with skin tone annotations by experts, crowds, and an algorithm. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–26, 2022. 3.4.2

[138] Consumer Reports Survey Group. *A.I./Algorithmic Decision-Making: Consumer Reports Nationally Representative Phone and Internet Survey, May 2024*. Consumer Reports, July 2024. URL https://advocacy.consumerreports.org/wp-content/uploads/2024/07/CR-AES-AI-Algorithms-Report-7.25.24.pdf. 3.4.1

[139] Marco Guermandi, Simone Benatti, Victor Javier Kartsch Morinigo, and Luca Bertini. A wearable device for minimally-invasive behind-the-ear eeg and evoked potentials. In *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2018. doi: 10.1109/BIOCAS.2018.8584814. 2.4.1 (a)

[140] Jisoo Ha, Seonghun Park, and Chang-Hwan Im. Novel hybrid brain-computer interface for virtual reality applications using steady-state visual-evoked potential-based brain–computer interface and electrooculogram-based eye tracking for increased information transfer rate. *Frontiers in Neuroinformatics*, 16, 2022. doi: 10.3389/fninf.2022.758537. 2.4.1 (a)

[141] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. Gender recognition or gender reductionism? the social implications of embedded gender recognition systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, page 1–13, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3173582. URL https://doi.org/10.1145/3173574.3173582. 3.7.4

[142] Anikó Hannák, Claudia Wagner, David Garcia, Alan Mislove, Markus Strohmaier, and Christo Wilson. Bias in online freelance marketplaces: Evidence from taskrabbit and fiverr. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1914–1933, 2017. 3.4.1

[143] Jane Hanson. AI is Replacing Humans In The Interview Process - What You Need To Know To Crush Your Next Video Interview. *Forbes*, September 2023. URL https://www.forbes.com/sites/janehanson/2023/09/30/ai-is-replacing-humans-in-the-interview-processwhat-you-need-to-know-to-crush-your-next-video-interview/. 3.3

[144] David Harborth and Alisa Frik. Evaluating and redefining smartphone permissions with contextualized justifications for mobile augmented reality apps. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 2021. ISBN 978-1-939133-25-0. URL https://www.usenix.org/conference/soups2021/presentation/harborth. 2.4.2, 2.9.5

[145] David Harborth and Sebastian Pape. Investigating privacy concerns related to mobile augmented reality apps–a vignette based online experiment. *Computers in Human Behavior*, 122, 2021. doi: 10.1016/j.chb.2021.106833. 2.4.3

[146] Scharon Harding. Lenovo announces consumer AR glasses that can tether to iPhones. *Ars Technica*, Sep 2022. URL https://arstechnica.com/gadgets/2022/09/lenovos-first-consumer-ar-glasses-to-debut-this-year-with-micro-oled-displays/. Accessed November 30, 2022. 2.4.1 (a)

[147] Elisa Harlan and Oliver Schnuck. Objective or biased: On the questionable use of artificial intelligence for job applications, 2021. URL https://interaktiv.br.de/ki-bewerbung/en/. 3.3, 3.7.4

[148] Christina N. Harrington and Lisa Egede. Trust, comfort and relatability: Understanding black older adults' perceptions of chatbot design for health information seeking. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. doi: 10.1145/3544548.3580719. URL https://doi.org/10.1145/3544548.3580719. 3.4.3

[149] Christina N. Harrington, Radhika Garg, Amanda Woodward, and Dimitri Williams. "It's Kind of Like Code-Switching": Black Older Adults' Experiences with a Voice Assistant for Health Information Seeking. In *CHI Conference on Human Factors in Computing Systems*, page 1–15, New Orleans LA USA, Apr 2022. ACM. ISBN 9781450391573. doi: 10.1145/3491102.3501995. URL https://dl.acm.org/doi/10.1145/3491102.3501995. 3.4.3

[150] Camille Harris, Amber Gayle Johnson, Sadie Palmer, Diyi Yang, and Amy Bruckman. "honestly, i think tiktok has a vendetta against black creators": Understanding black content creator experiences on tiktok. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2):1–31, 2023. 3.4.2

[151] Roy Harris. *The language myth*. Duckworth, London, 1981. ISBN 0715615289. 3.4.4

[152] Chris Harrison and Haakon Faste. Implications of location and touch for on-body projected interfaces. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, 2014. doi: 10.1145/2598510.2598587. 2.4.3

[153] Woodrow Hartzog. What is privacy? that's the wrong question. *The University of Chicago Law Review*, 88: 1677, 2021. 2.4.4

[154] Kirstie Hawkey, David Botta, Rodrigo Werlinger, Kasia Muldner, Andre Gagne, and Konstantin Beznosov. Human, Organizational, and Technological Factors of IT Security. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, pages 3639–3644. Association for Computing Machinery, 2008. ISBN 9781605580128. doi: 10.1145/1358628.1358905. URL https://doi.org/10.1145/1358628.1358905. 4.4.3

[155] Kathryn Henne, Renee Shelby, and Jenna Harb. The datafication of #metoo: Whiteness, racial capitalism, and anti-violence technologies. *Big Data & Society*, 8(2):20539517211055898, 2021. doi: 10.1177/20539517211055898. URL https://doi.org/10.1177/20539517211055898. 3.4.2

[156] Valentin Hofmann, Pratyusha Ria Kalluri, Dan Jurafsky, and Sharese King. AI generates covertly racist decisions about people based on their dialect. *Nature*, 633(8028):147–154, September 2024. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-024-07856-5. URL https://www.nature.com/articles/s41586-024-07856-5. 5.8.2

[157] Siegfried Hollerer, Thilo Sauter, and Wolfgang Kastner. Risk Assessments Considering Safety, Security, and Their Interdependencies in OT Environments. In *Proceedings of the 17th International Conference on Availability, Reliability and Security*, ARES '22, pages 1–8. Association for Computing Machinery, August 2022. ISBN 9781450396707. doi: 10.1145/3538969.3543814. URL https://dl.acm.org/doi/10.1145/3538969.3543814. 4.4.1, 4.4.2

[158] Siegfried Hollerer, Bernhard Brenner, Pushparaj Rajaram Bhosale, Clara Fischer, Ali Mohammad Hosseini, Sofia Maragkou, Maximilian Papa, Sebastian Schlund, Thilo Sauter, and Wolfgang Kastner. *Challenges in OT Security and Their Impacts on Safety-Related Cyber-Physical Production Systems*, pages 171–202. Springer, 2023. ISBN 9783662650042. doi: 10.1007/978-3-662-65004-2_7. URL https://doi.org/10.1007/978-3-662-65004-2_7. 4.4.1

[159] Faye Holt, William Held, and Diyi Yang. Perceptions of language technology failures from South Asian English speakers. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 4067–4081, Bangkok, Thailand, August 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.findings-acl.241. URL https://aclanthology.org/2024.findings-acl.241/. 3.4.3

[160] Jason Hong. Privacy and google glass, 2013. URL https://cacm.acm.org/blogs/blog-cacm/167230-privacy-and-google-glass/fulltext. Accessed November 30, 2022. 2.3

[161] The White House. Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, October 2023. URL https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/. 3.4.1

[162] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. Privacy behaviors of lifeloggers using wearable cameras. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2014. doi: 10.1145/2632048.2632079. 2.4.2 (c)

[163] Roberto Hoyle, Robert Templeman, Denise Anthony, David Crandall, and Apu Kapadia. Sensitive lifelogs: A privacy analysis of photos from wearable cameras. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015. doi: 10.1145/2702123.2702183. 2.4.2 (c)

[164] Elaine Hsieh. Understanding medical interpreters: Reconceptualizing bilingual health communication. *Health Communication*, 20(2):177–186, 2006. doi: 10.1207/s15327027hc2002_9. URL http://www.tandfonline.com/doi/abs/10.1207/s15327027hc2002_9. 5.3.3

[165] Elaine Hsieh. "I am not a robot!" Interpreters' Views of Their Roles in Health Care Settings. *Qualitative Health Research*, 18(10):1367–1383, 2008. doi: 10.1177/1049732308323840. URL https://journals.sagepub.com/doi/10.1177/1049732308323840. 5.3.3

[166] Elaine Hsieh. Not just "getting by": Factors influencing providers' choice of interpreters. *Journal of General Internal Medicine*, 30(1):75–82, January 2015. doi: 10.1007/s11606-014-3066-8. URL https://doi.org/10.1007/s11606-014-3066-8. 5.3.3

[167] Elaine Hsieh. *The Role of Healthcare Interpreters*, page 117–135. Wiley, 1 edition, March 2024. ISBN 9781119853817 9781119853855. doi: 10.1002/9781119853855.ch7. URL https://onlinelibrary.wiley.com/doi/10.1002/9781119853855.ch7. 5.3.3

[168] Jingting Huang, Yuhan Li, and Yiran Zheng. Gender discrimination in stem education. In *2022 6th International Seminar on Education, Management and Social Sciences (ISEMSS 2022)*, pages 1314–1323. Atlantis Press, 2022. 3.4.1

[169] Scott Hulver. What role do immigrants play in the hospital workforce? *KFF*, June 2025. URL https://www.kff.org/racial-equity-and-health-policy/what-role-do-immigrants-play-in-the-hospital-workforce/. 5.3.1

[170] Huma Imran, Mohamed Salama, Colin Turner, and Sherif Fattah. Cybersecurity Risk Management Frameworks in the Oil and Gas Sector: A Systematic Literature Review. In Kohei Arai, editor, *Advances in Information and Communication*, Lecture Notes in Networks and Systems, pages 871–894. Springer International Publishing, 2022. ISBN 9783030980153. doi: 10.1007/978-3-030-98015-3_59. 4.4.2

[171] ISC2 Inc. *ISC2 Cybersecurity Workforce Study 2023*. 2023. https://media.isc2.org/-/media/Project/ISC2/Main/Media/documents/research/ISC2_Cybersecurity_Workforce_Study_2023.pdf. 4.3

[172] Magic Leap Inc. Privacy policy, Jun 2020. URL https://resources.magicleap.com/en-us/privac

y/privacy-policy. Accessed November 30, 2022. 2.4.1 (b)

[173] Magic Leap Inc. Magic leap 2 full technology specifications, 2022. URL https://www.magicleap.com/magic-leap-2. Accessed November 30, 2022. 2.4.1 (b)

[174] Snap Inc. Privacy policy, Jun 2022. URL https://snap.com/en-US/privacy/privacy-policy. Accessed November 30, 2022. 2.4.1 (b)

[175] Snap Inc. Spectacles 3: Tech specs, 2022. URL https://www.spectacles.com/shop/spectacles-3. Accessed November 30, 2022. 2.4.1 (b)

[176] Immigration Research Initiative and Cyierra Roldan. Immigrants Are a Vital Part of America's Future - Immigration Research Initiative, October 2024. URL https://immresearch.org/publications/states/. 5.3.1

[177] Iulia Ion, Rob Reeder, and Sunny Consolvo. "...No one Can Hack My Mind": Comparing Expert and Non-Expert Security Practices. In *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, pages 327–346. USENIX Association, July 2015. ISBN 978-1-931971-249. URL https://www.usenix.org/conference/soups2015/proceedings/presentation/ion. 4.4.3

[178] Suhaila Ismail, Elena Sitnikova, and Jill Slay. Using integrated system theory approach to assess security for SCADA systems cyber security for critical infrastructures: A pilot study. In *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, pages 1000–1006, August 2014. doi: 10.1109/FSKD.2014.6980976. 4.4.2

[179] Elizabeth A. Jacobs, Donald S. Shepard, Jose A. Suaya, and Esta-Lee Stone. Overcoming language barriers in health care: Costs and benefits of interpreter services. *American Journal of Public Health*, 94(5):866–869, 2004. doi: 10.2105/AJPH.94.5.866. URL https://ajph.aphapublications.org/doi/full/10.2105/AJPH.94.5.866. 5.3.3

[180] Timo Jakobi, Maximilian Von Grafenstein, Patrick Smieskol, and Gunnar Stevens. A taxonomy of user-perceived privacy risks to foster accountability of data-based services. *Journal of Responsible Technology*, 10:100029, July 2022. doi: 10.1016/j.jrt.2022.100029. URL https://linkinghub.elsevier.com/retrieve/pii/S2666659622000063. 5.3.8

[181] Suman Jana, David Molnar, Alexander Moshchuk, Alan Dunn, Benjamin Livshits, Helen J. Wang, and Eyal Ofek. Enabling Fine-Grained Permissions for Augmented Reality Applications with Recognizers. In *22nd USENIX Security Symposium (USENIX Security 13)*, pages 415–430, Washington, D.C., August 2013. USENIX Association. ISBN 978-1-931971-03-4. URL https://www.usenix.org/conference/usenixsecurity13/technical-sessions/presentation/jana. 2.9.5

[182] Jan Jancar, Marcel Fourné, Daniel De Almeida Braga, Mohamed Sabt, Peter Schwabe, Gilles Barthe, Pierre-Alain Fouque, and Yasemin Acar. "They're not that hard to mitigate": What Cryptographic Library Developers Think About Timing Attacks. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 632–649, May 2022. doi: 10.1109/SP46214.2022.9833713. 4.4.3

[183] Scott Jaschik. Admissions Offices, Cautiously, Start Using AI, May 2023. URL https://www.insidehighered.com/news/admissions/2023/05/15/admissions-offices-cautiously-start-using-ai. 3.3

[184] Eun Seo Jo and Timnit Gebru. Lessons from archives: strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, page 306–316, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450369367. doi: 10.1145/3351095.3372829. URL https://doi.org/10.1145/3351095.3372829. 3.7.2

[185] Eunkyung Jo, Yuin Jeong, Sohyun Park, Daniel A. Epstein, and Young-Ho Kim. Understanding the impact of long-term memory on self-disclosure with large language model-driven chatbots for public health intervention. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, page 1–21, New York, NY, USA, May 2024. Association for Computing Machinery. ISBN 9798400703300. doi: 10.1145/3613904.3642420. URL https://dl.acm.org/doi/10.1145/3613904.3642420. 5.3.8

[186] Derek B. Johnson. Department of energy opens $9 million in competitive cyber funding to small electric

utilities. *SC Media*, August 2023. https://www.scmagazine.com/news/department-of-energy-opens-9-million-in-competitive-cyber-funding-to-small-electric-utilities. 4.3

[187] Pratik Joshi, Christain Barnes, Sebastin Santy, Simran Khanuja, Sanket Shah, Anirudh Srinivasan, Satwik Bhattamishra, Sunayana Sitaram, Monojit Choudhury, and Kalika Bali. Unsung challenges of building and deploying language technologies for low resource language communities, December 2019. URL http://arxiv.org/abs/1912.03457. arXiv:1912.03457 [cs]. 3.7.2

[188] Anjali Kantharuban, Ivan Vulić, and Anna Korhonen. Quantifying the dialect gap and its correlates across languages, October 2023. URL http://arxiv.org/abs/2310.15135. arXiv:2310.15135 [cs]. 3.7.2

[189] Jodi Kantor, Arya Sundaram, Aliza Aufrichtig, and Rumsey Taylor. The rise of the worker productivity score. *The New York Times*, August 2022. URL https://www.nytimes.com/interactive/2022/08/14/business/worker-productivity-tracking.html. 3.3

[190] Leah S. Karliner, Elizabeth A. Jacobs, Alice Hm Chen, and Sunita Mutha. Do professional interpreters improve clinical care for patients with limited english proficiency? a systematic review of the literature. *Health Services Research*, 42(2):727–754, 2007. doi: https://doi.org/10.1111/j.1475-6773.2006.00629.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6773.2006.00629.x. 5.3.3

[191] Joseph M. Kaufert and Robert W. Putsch. Communication through interpreters in healthcare: Ethical dilemmas arising from differences in class, culture, language, and power. *The Journal of Clinical Ethics*, 8(1):71–87, March 1997. doi: 10.1086/JCE199708111. URL https://www.journals.uchicago.edu/doi/10.1086/JCE199708111. 5.3.8

[192] Jasmeet Kaur, Preetika Sharma, Vijay Kumar, Mona Duggal, Nadia Griffin Diamond-Smith, Alison El Ayadi, Kathryn Vosburg, and Pushpendra Singh. Exploring the role of chatbots in tackling covid-19 vaccine hesitancy among pregnant and breastfeeding women in rural northern india. *Proc. ACM Hum.-Comput. Interact.*, 8 (CSCW1), April 2024. doi: 10.1145/3637332. URL https://doi.org/10.1145/3637332. 5.3.8

[193] Patrick Gage Kelley, Michael Benisch, Lorrie Faith Cranor, and Norman Sadeh. When are users comfortable sharing locations with advertisers? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2011. doi: 10.1145/1978942.1979299. 2.4.3

[194] Suleman Khan, M. Hammad Javed, Ehtasham Ahmed, Syed A A Shah, and Syed Umaid Ali. Facial recognition using convolutional neural networks and implementation on smart glasses. In *2019 International Conference on Information Science and Communication Technology (ICISCT)*, 2019. doi: 10.1109/CISCT.2019.8777442. 2.4.1 (a)

[195] Elaine C. Khoong and Alicia Fernandez. Addressing gaps in interpreter use: Time for implementation science informed multi-level interventions. *Journal of General Internal Medicine*, 36(11):3532–3536, 2021. doi: 10.1007/s11606-021-06823-4. URL https://link.springer.com/10.1007/s11606-021-06823-4. 5.3.3

[196] Elaine C Khoong and Jorge A Rodriguez. A research agenda for using machine translation in clinical medicine. *Journal of General Internal Medicine*, 37(5):1275–1277, 2022. 5.3.5

[197] Elaine C Khoong, Eric Steinbrook, Cortlyn Brown, and Alicia Fernandez. Assessing the use of google translate for spanish and chinese translations of emergency department discharge instructions. *JAMA internal medicine*, 179(4):580–582, 2019. 5.3.5

[198] Dhruv Khullar and Dave A Chokshi. Challenges for immigrant health in the usa—the road to crisis. *The Lancet*, 393(10186):2168–2174, 2019. doi: 10.1016/S0140-6736(19)30035-2. URL https://linkinghub.elsevier.com/retrieve/pii/S0140673619300352. 5.3.4

[199] Sara Kiesler and Lee S. Sproull. Response effects in the electronic survey. *Public Opinion Quarterly*, 50(3):402, 1986. doi: 10.1086/268992. URL https://academic.oup.com/poq/article-lookup/doi/10.1086/268992. 5.3.8

[200] Dokyun Kim, Wooseok Byun, Yunseo Ku, and Ji-Hoon Kim. High-speed visual target identification for low-cost wearable brain-computer interfaces. *IEEE Access*, 7, 2019. doi: 10.1109/ACCESS.2019.2912997. 2.4.1 (a)

[201] Yeji Kim. Virtual reality data and its privacy regulatory challenges: A call to move beyond text-based informed consent. *California Law Review*, 110(1), 2022. 2.9.5

[202] Sharese King and Katherine D. Kinzler. Op-Ed: Bias against African American English speakers is a pillar of systemic racism. *The Los Angeles Times*, July 2020. URL https://www.latimes.com/opinion/story/2020-07-14/african-american-english-racism-discrimination-speech. 3.4.4

[203] Sara Kingsley, Proteeti Sinha, Clara Wang, Motahhare Eslami, and Jason I. Hong. "Give Everybody [..] a Little Bit More Equity"': Content Creator Perspectives and Responses to the Algorithmic Demonetization of Content Associated with Disadvantaged Groups. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2), November 2022. doi: 10.1145/3555149. URL https://doi.org/10.1145/3555149. 3.4.2

[204] Kathryn T Knecht, Julie La, Kim-Phung Truong, Stacie Lo, Julia Chavez, Darlene F Tyler, and Paul Gavaza. Interactions of spanish-speaking latinas in a southern california community with their community pharmacists and pharmacy staff. *Journal of Contemporary Pharmacy Practice*, 69(3):19–26, 2022. doi: 10.37901/jcphp21-00010. URL https://meridian.allenpress.com/jcphp/article-pdf/69/3/19/3126686/i2573-2765-69-3-19.pdf. 5.3.3

[205] Liam Knox. Admissions offices deploy ai, October 2023. URL https://www.insidehighered.com/news/admissions/traditional-age/2023/10/09/admissions-offices-turn-ai-application-reviews. 3.3

[206] Marion Koelle, Matthias Kranz, and Andreas Möller. Don't look at me that way! understanding user attitudes towards data glasses usage. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 2015. doi: 10.1145/2785830.2785842. 2.4.2, 2.4.2 (a), 2.9.4

[207] Marion Koelle, Abdallah El Ali, Vanessa Cobus, Wilko Heuten, and Susanne CJ Boll. All about acceptability? identifying factors for the adoption of data glasses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017. doi: 10.1145/3025453.3025749. 2.4.2, 2.4.2 (a)

[208] Allison Koenecke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky, and Sharad Goel. Racial disparities in automated speech recognition. In *Proceedings of the National Academy of Sciences*, volume 117, page 7684–7689, Apr 2020. doi: 10.1073/pnas.1915768117. URL https://pnas.org/doi/full/10.1073/pnas.1915768117. 3.4.3, 5.3.6

[209] Dan Kosten. Immigrants as economic contributors: Immigrant tax contributions and spending power, September 2018. URL https://immigrationforum.org/article/immigrants-as-economic-contributors-immigrant-tax-contributions-and-spending-power/. 5.3.1

[210] Sara Kraemer and Pascale Carayon. A Human Factors Vulnerability Evaluation Method for Computer and Information Security. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47(12):1389–1393, 2003. doi: 10.1177/154193120304701202. URL http://journals.sagepub.com/doi/10.1177/154193120304701202. 4.4.3

[211] William A. Kretzschmar and Charles F. Meyer. *The idea of Standard American English*, page 139–158. Studies in English Language. Cambridge University Press, 2012. 3.3

[212] Jacob Leon Kröger, Otto Hans-Martin Lutz, and Florian Müller. What does your gaze reveal about you? on the privacy implications of eye tracking. In *IFIP International Summer School on Privacy and Identity Management*, 2019. doi: 10.1007/978-3-030-42504-3_15. 2.4.2

[213] Jacob Leon Kröger, Leon Gellrich, Sebastian Pape, Saba Rebecca Brause, and Stefan Ullrich. Personal information inference from voice recordings: User awareness and privacy concerns. *Proceedings on Privacy Enhancing Technologies*, 2022. 3.7.4

[214] Sahana Mukherjee Krogstad and Jens Manuel. Most U.S. voters say immigrants–no matter their legal status–mostly take jobs citizens don't want. *Pew Research Center*, October 2024. URL https://www.pewresearch.org/short-reads/2024/10/21/most-us-voters-say-immigrants-no-matter-their-legal-status-mostly-take-jobs-citizens-dont-want/. 5.3.1

[215] Leighton Ku and Glenn Flores. Pay now or pay later: Providing interpreter services in health care. *Health Affairs*, 24(2):435–444, 2005. doi: 10.1377/hlthaff.24.2.435. URL http://www.healthaffairs.org/do

i/10.1377/hlthaff.24.2.435. 5.3.3

[216] Jangho Kwon, Jihyeon Ha, Da-Hye Kim, Jun Won Choi, and Laehyun Kim. Emotion recognition using a glasses-type wearable device via multi-channel facial responses. *IEEE Access*, 9, 2021. doi: 10.1109/ACCESS.2021.3121543. 2.4.1 (a)

[217] William Labov. *A Study of Non-Standard English*. ERIC Clearinghouse for Languages and Linguistics, January 1969. URL https://eric.ed.gov/?id=ED024053. ERIC Number: ED024053. 3.3, 3.4.4

[218] William Labov. *The three dialects of English*, page 1–44. Quantitative analyses of linguistic structure. Academic Press, San Diego, 1991. ISBN 0122297903. 3.4.4

[219] William Labov, Sharon Ash, and Charles Boberg. A national map of the regional dialects of american english. *Linguistics Laboratory, Dept. of Linguistics, U of Pennsylvania*, 15, 1997. 3.4.4, 3.5.1

[220] Lutz Lammerding, Tim Hilken, Dominik Mahr, and Jonas Heller. Too real for comfort: Measuring consumers' augmented reality information privacy concerns. In *Augmented Reality and Virtual Reality: New Trends in Immersive Technology*, pages 95–108. Springer, 2021. 2.4.3

[221] Josephine Lamp, Carlos E. Rubio-Medrano, Ziming Zhao, and Gail-Joon Ahn. ExSol: Collaboratively Assessing Cybersecurity Risks for Protecting Energy Delivery Systems. In *2019 7th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, pages 1–6, April 2019. doi: 10.1109/MSCPES.2019.8738791. 4.4.2

[222] Katharine Lawrence, Stella K. Chong, Holly Krelle, Timothy R. Roberts, Lorna E Thorpe, Chau Trinh-Shevrin, Stella S Yi, and Simona C. Kwon. Chinese americans' use of patient portal systems: a scoping review (preprint). *JMIR human factors*, 2022. doi: 10.2196/27924. 5.3.3

[223] Roslyn Layton. Hackers Are Targeting U.S. Banks, And Hardware May Give Them An Open Door. *Forbes*, February 2021. https://www.forbes.com/sites/roslynlayton/2021/03/17/hackers-are-targeting-us-banks-and-hardware-may-give-them-an-open-door/. 4.3

[224] Kiron Lebeck, Kimberly Ruth, Tadayoshi Kohno, and Franziska Roesner. Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)*, 2018. doi: 10.1109/SP.2018.00051. 2.4.2, 2.4.2 (a), 2.9.4

[225] Hao-Ping (Hank) Lee, Yu-Ju Yang, Thomas Serban Von Davier, Jodi Forlizzi, and Sauvik Das. Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. doi: 10.1145/3613904.3642116. URL https://doi.org/10.1145/3613904.3642116. 3.7.4, 5.3.8

[226] Daniel Leufer. Privacy review of facebook's ray-ban stories smart glasses. *Access Now*, Sep 2021. URL https://www.accessnow.org/facebook-ray-ban-stories-smart-glasses-privacy-review/. Accessed November 30, 2022. 2.9.5

[227] James Andrew Lewis and William Crumpler. The cybersecurity workforce gap. *Center for Strategic and International Studies*, January 2019. https://www.csis.org/analysis/cybersecurity-workforce-gap. 4.3

[228] Huining Li, Chenhan Xu, Aditya Singh Rathore, Zhengxiong Li, Hanbin Zhang, Chen Song, Kun Wang, Lu Su, Feng Lin, Kui Ren, and Wenyao Xu. Vocalprint: Exploring a resilient and secure voice authentication via mmwave biometric interrogation. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*, 2020. doi: 10.1145/3384419.3430779. 2.4.1 (a)

[229] Jingquan Li. Security Implications of AI Chatbots in Health Care. *Journal of Medical Internet Research*, 25: e47551, November 2023. doi: 10.2196/47551. URL https://www.jmir.org/2023/1/e47551. 5.3.8

[230] Kai Li, Zhangxi Lin, and Xiaowen Wang. An empirical analysis of users' privacy disclosure behaviors on social network sites. *Information & management*, 52(7):882–891, 2015. doi: 10.1016/j.im.2015.07.006. 2.4.2 (b)

[231] Lan Li, Tina Lassiter, Joohee Oh, and Min Kyung Lee. Algorithmic hiring in practice: Recruiter and hr profes-

sional's perspectives on AI use in hiring. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '21, page 166–176, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450384735. doi: 10.1145/3461702.3462531. URL https://doi.org/10.1145/3461702.3462531. 3.4.1

[232] Q. Vera Liao and Jennifer Wortman Vaughan. AI transparency in the age of llms: A human-centered research roadmap. *Harvard Data Science Review*, (Special Issue 5), May 2024. doi: 10.1162/99608f92.8036d03b. URL https://hdsr.mitpress.mit.edu/pub/aelql9qy/release/2. 3.7.2

[233] Martin C. Libicki, David Senty, and Julia Pollak. *Hackers Wanted: An Examination of the Cybersecurity Labor Market*. June 2014. https://www.rand.org/pubs/research_reports/RR430.html. 4.3

[234] Library of Congress. ISO 639 Frequently Asked Questions (FAQ). https://www.loc.gov/standards/iso639-2/faq.html, 2023. Accessed on 2025-09-25. 5.4.4

[235] Library of Congress. Codes for the Representation of Names of Languages Part 2: Alpha-3 Code. https://www.loc.gov/standards/iso639-2/langhome.html, 2024. Updated on July 19, 2024; Accessed on 2025-09-25. 5.4.4

[236] Daniel J. Liebling and Sören Preibusch. Privacy considerations for a pervasive eye tracking world. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, 2014. doi: 10.1145/2638728.2641688. 2.4.2

[237] Daniel J. Liebling, Michal Lahav, Abigail Evans, Aaron Donsbach, Jess Holbrook, Boris Smus, and Lindsey Boran. Unmet needs and opportunities for mobile translation ai. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, page 1–13, Honolulu HI USA, April 2020. ACM. ISBN 9781450367080. doi: 10.1145/3313831.3376261. URL https://dl.acm.org/doi/10.1145/3313831.3376261. 5.2

[238] Fangru Lin, Shaoguang Mao, Emanuele La Malfa, Valentin Hofmann, Adrian de Wynter, Xun Wang, Si-Qing Chen, Michael J. Wooldridge, Janet B. Pierrehumbert, and Furu Wei. Assessing dialect fairness and robustness of large language models in reasoning tasks. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6317–6342, Vienna, Austria, July 2025. Association for Computational Linguistics. ISBN 979-8-89176-251-0. doi: 10.18653/v1/2025.acl-long.317. URL https://aclanthology.org/2025.acl-long.317/. 5.8.2

[239] Maria B. Line, Ali Zand, Gianluca Stringhini, and Richard Kemmerer. Targeted Attacks against Industrial Control Systems: Is the Power Industry Prepared? In *Proceedings of the 2nd Workshop on Smart Energy Grid Security*, SEGS '14, pages 13–22. Association for Computing Machinery, November 2014. ISBN 9781450331548. doi: 10.1145/2667190.2667192. URL https://doi.org/10.1145/2667190.2667192. 4.4.3

[240] P. Litherland, R. Orr, and R. Piggin. Cyber security of operational technology: understanding differences and achieving balance between nuclear safety and nuclear security. In *11th International Conference on System Safety and Cyber-Security (SSCS 2016)*, pages 1–6, January 2016. doi: 10.1049/cp.2016.0856. URL https://doi-org.cmu.idm.oclc.org/10.1049/cp.2016.0856. 4.4.1

[241] Lloyd's. Business blackout: The insurance implications of a cyber attack on the US power grid. https://www.lloyds.com/news-and-insights/risk-reports/library/business-blackout/, July 2015. 4.3

[242] Craig Locatis, Deborah Williamson, Carrie Gould-Kabler, Laurie Zone-Smith, Isabel Detzler, Jason Roberson, Richard Maisiak, and Michael Ackerman. Comparing in-person, video, and telephonic medical interpretation. *Journal of General Internal Medicine*, 25(4):345–350, 2010. doi: 10.1007/s11606-009-1236-x. URL http://link.springer.com/10.1007/s11606-009-1236-x. 5.3.3

[243] Nina Lutz and Cecilia Aragon. "We're not all construction workers": Algorithmic Compression of Latinidad on TikTok. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW2), November 2024. doi: 10.1145/3687019. URL https://doi.org/10.1145/3687019. 3.4.2

[244] C. Lyles, J. Fruchterman, M. Youdelman, and D. Schillinger. Legal, practical, and ethical considerations for making online patient portals accessible for all. *American journal of public health*, 2017. doi: 10.2105/ajph.2017.303933. 5.3.3

[245] Xiao Ma, Jeff Hancock, and Mor Naaman. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016. doi: 10.1145/2858036.2858414. 2.4.3

[246] Juan F. Maestre, Daria V. Groves, Megan Furness, and Patrick C. Shih. "it's like with the pregnancy tests": Co-design of speculative technology for public hiv-related stigma and its implications for social media. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, page 1–21, Hamburg Germany, April 2023. ACM. ISBN 9781450394215. doi: 10.1145/3544548.3581033. URL https://dl.acm.org/doi/10.1145/3544548.3581033. 5.3.8

[247] Angus Main and Dylan Yamada-Rice. Evading big brother: Using visual methods to understand children's perception of sensors and interest in subverting digital surveillance. *Visual Communication*, 21(3), 2022. doi: 10.1177/14703572221093559. 2.4.1 (a)

[248] Lev Malevanchik, Margaret Wheeler, Kristin Gagliardi, Leah Karliner, and Sachin J. Shah. Disparities after discharge: The association of limited english proficiency and post-discharge patient reported issues. *Joint Commission Journal on Quality and Patient Safety*, 47(12):775–782, December 2021. doi: 10.1016/j.jcjq.2021.08.013. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9246478/. 5.3.3

[249] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. Anonymity vs. familiarity: Self-disclosure and privacy in social virtual reality. In *26th ACM Symposium on Virtual Reality Software and Technology*, 2020. doi: 10.1145/3385956.3418967. 2.4.2 (b)

[250] Alessandro Mantovani, Simone Aonzo, Yanick Fratantonio, and Davide Balzarotti. RE-Mind: a First Look Inside the Mind of a Reverse Engineer. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 2727–2745. USENIX Association, August 2022. ISBN 978-1-939133-31-1. URL https://www.usenix.org/conference/usenixsecurity22/presentation/mantovani. 4.4.3

[251] Nina Markl. Language variation and algorithmic bias: understanding algorithmic bias in British English automatic speech recognition. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, page 521–534, New York, NY, USA, June 2022. Association for Computing Machinery. ISBN 9781450393522. doi: 10.1145/3531146.3533117. URL https://doi.org/10.1145/3531146.3533117. 3.3, 3.4.3, 5.3.6

[252] Nina Markl and Catherine Lai. Context-sensitive evaluation of automatic speech recognition: considering user experience & language variation. In *Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing*, page 34–40, Online, April 2021. Association for Computational Linguistics. URL https://aclanthology.org/2021.hcinlp-1.6. 3.4.3

[253] Nina Markl and Stephen Joseph McNulty. Language technology practitioners as language managers: arbitrating data bias and predictive bias in asr. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, page 6328–6339, Marseille, France, June 2022. European Language Resources Association. URL https://aclanthology.org/2022.lrec-1.680. 3.4.3

[254] Nina Markl, Electra Wallington, Ondrej Klejch, Thomas Reitmaier, Gavin Bailey, Jennifer Pearson, Matt Jones, Simon Robinson, and Peter Bell. Automatic transcription and (de)standardisation. In *Proceedings - SIGUL 2023, 2nd Annual Meeting of the Special Interest Group on Under-resourced Languages: A Satellite Workshop of Interspeech 2023*, page 93–97. International Speech Communication Association, August 2023. doi: 10.21437/SIGUL.2023-20. URL https://www.research.ed.ac.uk/en/publications/automatic-transcription-and-destandardisation. 3.4.3

[255] Mason Marks and Claudia E. Haupt. AI Chatbots, Health Privacy, and Challenges to HIPAA Compliance. *JAMA*, 330(4):309, July 2023. doi: 10.1001/jama.2023.9458. URL https://jamanetwork.com/journals/jama/fullarticle/2807170. 5.3.8

[256] Claire Marshall and Malcom Prior. Cyber security: Global food supply chain at risk from malicious hackers.

*BBC News*, May 2022. https://www.bbc.com/news/science-environment-61336659. 4.3

[257] Joshua L. Martin and Kevin Tang. Understanding Racial Disparities in Automatic Speech Recognition: The Case of Habitual "be". In *Proceedings of Interspeech 2020*, page 626–630, 2020. doi: 10.21437/Interspeech.2020-2893. URL https://www.isca-archive.org/interspeech_2020/martin20_interspeech.html. 3.4.3

[258] Alice E Marwick and Danah Boyd. Networked privacy: How teenagers negotiate context in social media. *New media & society*, 16(7):1051–1067, 2014. doi: 10.1177/146144481454399. 2.9.2

[259] Katsutoshi Masai, Yuta Sugiura, Masa Ogata, Kai Kunze, Masahiko Inami, and Maki Sugimoto. Facial expression recognition in daily life by embedded photo reflective sensors on smart eyewear. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 2016. doi: 10.1145/2856767.2856770. 2.4.1 (a)

[260] Mary C. Masland, Christine Lou, and Lonnie Snowden. Use of communication technologies to cost-effectively increase the availability of interpretation services in healthcare settings. *Telemedicine and e-Health*, 16(6):739–745, 2010. doi: 10.1089/tmj.2009.0186. URL https://www.liebertpub.com/doi/10.1089/tmj.2009.0186. 5.3.3

[261] Richard May and Kerstin Denecke. Security, privacy, and healthcare-related conversational agents: a scoping review. *Informatics for Health and Social Care*, 47(2):194–210, April 2022. doi: 10.1080/17538157.2021.1983578. URL https://www.tandfonline.com/doi/full/10.1080/17538157.2021.1983578. 5.3.8

[262] Nora McDonald and Andrea Forte. The politics of privacy theories: Moving from norms to vulnerabilities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020. doi: 10.1145/3313831.3376167. 2.4.2 (b)

[263] Lavinia McLean and Mark D Griffiths. Female gamers' experience of online harassment and social support in online gaming: A qualitative study. *International Journal of Mental Health and Addiction*, 17(4):970–994, 2019. doi: 10.1007/s11469-018-9962-0. 2.4.2 (b)

[264] Nikita Mehandru, Samantha Robertson, and Niloufar Salehi. Reliable and safe use of machine translation in medical settings. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, page 2016–2025, New York, NY, USA, June 2022. Association for Computing Machinery. ISBN 9781450393522. doi: 10.1145/3531146.3533244. URL https://dl.acm.org/doi/10.1145/3531146.3533244. 5.3.5

[265] Tamir Mendel, Oded Nov, and Batia Wiesenfeld. Advice from a doctor or ai? understanding willingness to disclose information through remote patient monitoring to receive health advice. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW2), November 2024. doi: 10.1145/3686925. URL https://doi.org/10.1145/3686925. 5.3.8

[266] Meta. Hand tracking privacy notice, 2022. URL https://store.facebook.com/en-gb/help/quest/articles/accounts/privacy-information-and-settings/hand-tracking-privacy-notice/. Accessed November 30, 2022. 2.4.1 (b)

[267] Meta. Oculus privacy policy | meta quest, 2022. URL https://store.facebook.com/de/en/legal/quest/updated-privacy-policy-for-oculus-account-users/. Accessed November 30, 2022. 2.4.1 (b)

[268] Meta. Monitoring and recording safety horizon | meta quest, 2022. URL https://store.facebook.com/de/en/legal/quest/monitoring-recording-safety-horizon/. Accessed November 30, 2022. 2.4.1 (b)

[269] Meta. Privacy information and settings, 2022. URL https://store.facebook.com/en-gb/help/quest/articles/accounts/privacy-information-and-settings/. Accessed November 30, 2022. 2.4.1 (b)

[270] Cade Metz. Can a.i. grade your next test? *The New York Times*, July 2021. URL https://www.nytimes.com/2021/07/20/technology/ai-education-neural-networks.html. 3.3

[271] Rachel Metz. There's a new obstacle to landing a job after college: Getting approved by AI | cnn business, January 2020. URL https://www.cnn.com/2020/01/15/tech/ai-job-interview. 3.3

[272] Barbara Micale. AI algorithms can add objectivity in assessing job applicants, March 2024. URL https://news.vt.edu/content/news_vt_edu/en/articles/2024/02/science-hickman-ai-algorithms.html. 3.3

[273] Ola Michalec, Sveta Milyaeva, and Awais Rashid. When the future meets the past: Can safety and cyber security coexist in modern critical infrastructures? *Big Data & Society*, 9(1):205395172211083, 2022. doi: 10.1177/20539517221108369. URL http://journals.sagepub.com/doi/10.1177/20539517221108369. 4.3, 4.4.1, 4.7.1

[274] Ola Aleksandra Michalec, Dirk van der Linden, Sveta Milyaeva, and Awais Rashid. Industry Responses to the European Directive on Security of Network and Information Systems (NIS): Understanding policy implementation practices across critical infrastructures. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, pages 301–317. USENIX Association, August 2020. ISBN 978-1-939133-16-8. URL https://www.usenix.org/conference/soups2020/presentation/michalec. 4.4.3

[275] Microsoft. Improve visual quality and comfort, Sep 2022. URL https://learn.microsoft.com/en-us/hololens/hololens-calibration. Accessed November 30, 2022. 2.4.1 (a)

[276] Microsoft. Hololens 2 hardware, Jul 2022. URL https://docs.microsoft.com/en-us/hololens/hololens2-hardware. Accessed November 30, 2022. 2.4.1 (a)

[277] Claire Cain Miller. Can an Algorithm Hire Better Than a Human? *The New York Times*, June 2015. URL https://www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html. 3.3

[278] Maggie Miller. The mounting death toll of hospital cyberattacks. *POLITICO*, December 2022. https://www.politico.com/news/2022/12/28/cyberattacks-u-s-hospitals-00075638. 4.3

[279] Jaron Mink, Harjot Kaur, Juliane Schmüser, Sascha Fahl, and Yasemin Acar. "Security is not my field, I'm a stats guy": A Qualitative Root Cause Analysis of Barriers to Adversarial Machine Learning Defenses in Industry. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 3763–3780. USENIX Association, August 2023. ISBN 978-1-939133-37-3. URL https://www.usenix.org/conference/usenixsecurity23/presentation/mink. 4.4.3

[280] Tom Moberly. Doctors are cautioned against using google translate in consultations. *BMJ*, 363, 2018. ISSN 0959-8138. doi: 10.1136/bmj.k4546. URL https://www.bmj.com/content/363/bmj.k4546. 5.3.2, 5.3.5

[281] Penn Wharton Budget Model. *The Effects of Immigration on the United States' Economy*. Penn Wharton Budget Model, June 2016. URL https://budgetmodel.wharton.upenn.edu/issues/2016/1/27/the-effects-of-immigration-on-the-united-states-economy. 5.3.1

[282] Glenn Murray, Michael N. Johnstone, and Craig Valli. The convergence of IT and OT in critical infrastructure. In *Proceedings of the 15th Australian Information Security Management Conference*, pages 149–155, December 2017. doi: 10.4225/75/5a84f7b595b4e. 4.4.1

[283] Anish Nag, Nick Haber, Catalin Voss, Serena Tamura, Jena Daniels, Jeffrey Ma, Bryan Chiang, Shasta Ramachandran, Jessey Schwartz, Terry Winograd, Carl Feinstein, and Dennis P. Wall. Toward continuous social phenotyping: Analyzing gaze patterns in an emotion recognition task for children with autism through wearable smart glasses. *Journal of Medical Internet Research*, 22(4), 2020. doi: 10.2196/13810. 2.4.1 (a)

[284] Cen Nan and Irene Eusgeld. Exploring impacts of single failure propagation between SCADA and SUC. In *2011 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 1564–1568, December 2011. doi: 10.1109/IEEM.2011.6118180. 4.4.2

[285] Alondra Nelson. Three Fallacies: Alondra Nelson's Remarks at the Elysée Palace on the Occasion of the AI Action Summit | TechPolicy.Press. *Tech Policy Press*, February 2025. URL https://techpolicy.press/three-fallacies-alondra-nelsons-remarks-at-elyse-palace-on-the-occasion-of-the-ai-action-summit. 6.3.2

[286] Nicholas V Nguyen, Andres H Guillen Lozoya, Maria A Caruso, Maria Graciela D Capetillo Porraz, Laura M Pacheco-Spann, Megan A Allyse, and Amelia K Barwise. Through the eyes of spanish-speaking patients, caregivers, and community leaders: a qualitative study on the in-patient hospital experience. *International Journal for Equity in Health*, 23(1):164, 2024. 5.3.3

[287] Hellina Hailu Nigatu, John Canny, and Sarah E. Chasins. Low-resourced languages and online knowledge repositories: A need-finding study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, page 1–21, New York, NY, USA, May 2024. Association for Computing Machinery. ISBN 9798400703300. doi: 10.1145/3613904.3642605. URL https://dl.acm.org/doi/10.1145/3613904.3642605. 3.7.2

[288] Helen Nissenbaum. Toward an approach to privacy in public: Challenges of information technology. *Ethics & Behavior*, 7(3):207–219, 1997. doi: 10.1207/s15327019eb0703_3. 2.4.4

[289] Helen Nissenbaum. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford Law Books, Stanford, California, 2010. ISBN 9780804772891. 2.4.4, 2.9.2

[290] Helen Nissenbaum. A Contextual Approach to Privacy Online. *Daedalus*, 140(4), 2011. doi: 10.1162/DAED_a_00113. 2.4.4, 2.9.2

[291] North American Electric Reliability Corporation. Reliability standards. https://www.nerc.com/pa/Stand/Pages/ReliabilityStandards.aspx, 2023. Accessed February 22, 2024. 4.4.2, 4.6.1 (a)

[292] Fayika Farhat Nova, Michael Ann DeVito, Pratyasha Saha, Kazi Shohanur Rashid, Shashwata Roy Turzo, Sadia Afrin, and Shion Guha. "facebook promotes more harassment": Social media ecosystem, skill and marginalized hijra identity in bangladesh. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CSCW1), Apr 2021. doi: 10.1145/3449231. URL https://doi.org/10.1145/3449231. 2.4.2 (b)

[293] Iii Alejandro Ochoa, K. Kitayama, S. Uijtdehaage, M. Vermillion, Michael Eaton, Felix Carpio, Martin Serota, and M. Hochman. Patient and provider perspectives on the potential value and use of a bilingual online patient portal in a spanish-speaking safety-net population. *J. Am. Medical Informatics Assoc.*, 2017. doi: 10.1093/jamia/ocx040. 5.3.3

[294] Danielle Ochs, Jennifer Betts, and Zachary V. Zagger. California's wait is nearly over: New AI employment discrimination regulations move toward final publication. *Ogletree*, April 2025. URL https://ogletree.com/insights-resources/blog-posts/californias-wait-is-nearly-over-new-ai-employment-discrimination-regulations-move-toward-final-publication/. 3.7.4

[295] Federation of American Scientists Intelligence Resource Program. FM 34-36 Appendix D: Target Analysis Practice. https://irp.fas.org/doddir/army/fm34-36/appd.htm, 1991. Accessed February 22, 2024. 4.6.1 (a)

[296] Office of Educational Technology. Artificial intelligence and the future of teaching and learning. Technical report, Office of Educational Technology, May 2023. URL https://tech.ed.gov/files/2023/05/ai-future-of-teaching-and-learning-report.pdf. 3.4.1

[297] The Department of Health and Human Services (HHS). Limited english proficiency (lep). URL https://www.hhs.gov/civil-rights/for-individuals/special-topics/limited-english-proficiency/index.html. 5.3.2

[298] Forum of Incident Response and Security Teams. CVSS v3.1 Specification Document. https://www.first.org/cvss/specification-document, 2023. Accessed February 22, 2024. 4.4.2, 4.6.1 (a)

[299] U.S. Bureau of Labor Statistics. *In 2023, the majority of home health aides and personal care aides were women*. TED: The Economics Daily. U.S. Bureau of Labor Statistics, November 2024. URL https://www.bls.gov/opub/ted/2024/in-2023-the-majority-of-home-health-aides-and-personal-care-aides-were-women.htm. 5.3.1

[300] U.S. Bureau of Labor Statistics. *Labor Force Characteristics of Foreign-born Workers News Release - 2024 A01 Results*. Number USDL-25-0847. U.S. Bureau of Labor Statistics, May 2025. URL https://www.bls.gov/news.release/forbrn.htm. 5.3.1

[301] National Institute of Standards and Technology. NVD - CVEs and the NVD Process. https://nvd.nist.gov/general/cve-process. Accessed February 22, 2024. 4.6.1 (a)

[302] National Institute of Standards and Technology. Cybersecurity Framework | NIST. https://www.nist.gov/cyberframework, April 2018. Accessed February 22, 2024. 4.6.1 (a)

[303] National Institute of Standards and Technology. Vulnerability - Glossary CSRC. https://csrc.nist.gov/glossary/term/vulnerability, 2019. Accessed February 22, 2024. 4.3

[304] National Institute of Standards and Technology. NVD - Vulnerability Metrics. https://nvd.nist.gov/vuln-metrics, 2023. Accessed February 22, 2024. 4.4.2, 4.6.1 (a)

[305] U. S. Government Accountability Office. Preparing for Evolving Cybersecurity Threats Facing the U.S. Electric Grid. https://www.gao.gov/blog/2019/10/16/preparing-for-evolving-cybersecurity-threats-facing-the-u-s-electric-grid, Sep 2019. 4.3

[306] Joseph O'Hagan, Pejman Saeghe, Jan Gugenheimer, Daniel Medeiros, Karola Marky, Mohamed Khamis, and Mark McGill. Privacy-enhancing technology and everyday augmented reality: Understanding bystanders' varying needs for awareness and consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 6(4), Jan 2023. doi: 10.1145/3569501. URL https://doi.org/10.1145/3569501. 2.4.2 (a), 2.9.4

[307] Nicole L. Olenik, Jasmine D. Gonzalvo, Margie E. Snyder, Christy L. Nash, and Cory T. Smith. Perceptions of spanish-speaking clientele of patient care services in a community pharmacy. *Research in Social & Administrative Pharmacy: RSAP*, 11(2):241–252, 2015. doi: 10.1016/j.sapharm.2014.07.001. 5.3.3

[308] Jon Oltsik and Bill Lundell. *The Life and Times of Cybersecurity Professionals 2021 Volume V*. 2021. https://www.issa.org/wp-content/uploads/2021/07/ESG-ISSA-Research-Report-Life-of-Cybersecurity-Professionals-Jul-2021.pdf. 4.3

[309] Wanda J. Orlikowski. Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science*, 11(4):404–428, 2000. doi: 10.1287/orsc.11.4.404.14600. URL https://doi.org/10.1287/orsc.11.4.404.14600. 3.4.2

[310] Pilar Ortega, Mónica Vela, and Elizabeth A. Jacobs. Raising the bar for language equity health care research. *JAMA Network Open*, 6(7):e2324485, July 2023. doi: 10.1001/jamanetworkopen.2023.24485. URL https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2807144. 5.3.4

[311] Raphael Ouzan. AI broke hiring. can generative AI fix it? *Fast Company*, August 2024. URL https://www.fastcompany.com/91173678/ai-broke-hiring-can-generative-ai-fix-it. 3.3

[312] Anita Panayiotou, Anastasia Gardner, Sue Williams, Emiliano Zucchi, Monita Mascitti-Meuter, Anita MY Goh, Emily You, Terence WH Chong, Dina Logiudice, Xiaoping Lin, Betty Haralambous, and Frances Batchelor. Language Translation Apps in Health Care Settings: Expert Opinion. *JMIR mHealth and uHealth*, 7 (4):e11316, April 2019. doi: 10.2196/11316. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6477569/. 5.3.5

[313] Orestis Papakyriakopoulos and Alice Xiang. Considerations for ethical speech recognition datasets. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, WSDM '23, page 1287–1288, New York, NY, USA, Feb 2023. Association for Computing Machinery. ISBN 9781450394079. doi: 10.1145/3539597.3575793. URL https://doi.org/10.1145/3539597.3575793. 3.4.3

[314] Orestis Papakyriakopoulos, Anna Seo Gyeong Choi, William Thong, Dora Zhao, Jerone Andrews, Rebecca Bourke, Alice Xiang, and Allison Koenecke. Augmented datasheets for speech datasets and ethical decision-making. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '23, page 881–904, New York, NY, USA, Jun 2023. Association for Computing Machinery. ISBN 9798400701924. doi: 10.1145/3593013.3594049. URL https://dl.acm.org/doi/10.1145/3593013.3594049. 3.4.3

[315] Chrysanthi Papoutsi and Ian Brown. Privacy as articulation work in hiv health services. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, page 339–348, New York, NY, USA, February 2015. Association for Computing Machinery. ISBN 9781450329224.

doi: 10.1145/2675133.2675204. URL https://dl.acm.org/doi/10.1145/2675133.2675204. 5.3.8

[316] Priya R Pathak, Melissa S Stockwell, Mariellen M Lane, Laura Robbins-Milne, Suzanne Friedman, Kalpana Pethe, Margaret C Krause, Karen Soren, Luz Adriana Matiz, Lauren B Solomon, Maria E Burke, and Edith Bracho-Sanchez. Access to primary care telemedicine and visit characterization in a pediatric, low-income, primarily latino population: Retrospective study. *JMIR Pediatrics and Parenting*, 7:e57702, December 2024. doi: 10.2196/57702. URL https://pediatrics.jmir.org/2024/1/e57702. 5.3.4

[317] Destiny Peterson. In the Thick of Thick Accents: Employment Discrimination and the Appalachian Accent. *Appalachian Journal of Law*, 22(2), June 2023. URL https://appalachian.scholasticahq.com/article/74188-in-the-thick-of-thick-accents-employment-discrimination-and-the-appalachian-accent. 3.4.4

[318] Thao Phan, Jake Goldenfein, Monique Mann, and Declan Kuch. Economies of Virtue: The Circulation of 'Ethics' in Big Tech. *Science as Culture*, 31(1), 2022. doi: 10.1080/09505431.2021.1990875. 2.9.5

[319] Fisher Phillips. California regulators adopt new discrimination rules for automated-decision systems: 3 steps for employers using AI in the workplace. *JD Supra*, April 2025. URL https://www.jdsupra.com/legalnews/california-regulators-adopt-new-9904630/. 3.7.4

[320] Lindsey M. Philpot, Priya Ramar, Daniel L. Roellinger, Margaret A. McIntee, and Jon O. Ebbert. Digital health literacy and use of patient portals among spanish-preferred patients in the united states: a cross-sectional assessment. *Frontiers in Public Health*, Volume 12 - 2024, 2024. ISSN 2296-2565. doi: 10.3389/fpubh.2024.1455395. URL https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2024.1455395. 5.3.7

[321] Claudio Pinhanez, Paulo Cavalin, Luciana Storto, Thomas Finbow, Alexander Cobbinah, Julio Nogima, Marisa Vasconcelos, Pedro Domingues, Priscila de Souza Mizukami, Nicole Grell, Majoí Gongora, and Isabel Gonçalves. Harnessing the power of artificial intelligence to vitalize endangered indigenous languages: Technologies and experiences, July 2024. URL http://arxiv.org/abs/2407.12620. 3.7.2

[322] Claudio Santos Pinhanez, Raul Fernandez, Marcelo Carpinette Grave, Julio Nogima, and Ron Hoory. Creating an African American-Sounding TTS: Guidelines, Technical Challenges, and Surprising Evaluations. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, page 259–273, Greenville SC USA, March 2024. ACM. ISBN 9798400705083. doi: 10.1145/3640543.3645165. URL https://dl.acm.org/doi/10.1145/3640543.3645165. 3.7.2

[323] Blaine A. Price, Avelie Stuart, Gul Calikli, Ciaran Mccormick, Vikram Mehta, Luke Hutton, Arosha K. Bandara, Mark Levine, and Bashar Nuseibeh. Logging you, logging me: A replicable study of privacy and sharing behaviour in groups of visual lifeloggers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2), 2017. doi: 10.1145/3090087. 2.4.2 (c)

[324] Erika Leemann Price, Eliseo J. Pérez-Stable, Dana Nickleach, Monica López, and Leah S. Karliner. Interpreter perspectives of in-person, telephonic, and videoconferencing medical interpretation in clinical encounters. *Patient Education and Counseling*, 87(2):226–232, 2012. doi: 10.1016/j.pec.2011.08.006. URL https://linkinghub.elsevier.com/retrieve/pii/S0738399111004551. 5.3.3

[325] Kerri Prinos, Neal Patwari, and Cathleen A. Power. Speaking of accent: A content analysis of accent misconceptions in asr research. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, page 1245–1254, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704505. doi: 10.1145/3630106.3658969. URL https://doi.org/10.1145/3630106.3658969. 3.4.3, 5.3.6

[326] Jason Procyk, Carman Neustaedter, Carolyn Pang, Anthony Tang, and Tejinder K. Judge. Exploring video streaming in public settings: Shared geocaching over distance using mobile video chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014. doi: 10.1145/2556288.2557198. 2.4.2 (c)

[327] Thomas Purnell, William Idsardi, and John Baugh. Perceptual and phonetic experiments on american english dialect identification. *Journal of Language and Social Psychology*, 18(1):10–30, 1999. doi: 10.1177/0261

927X99018001002. URL https://journals.sagepub.com/doi/10.1177/0261927X99018001002. 3.4.4

[328] Cassidy Pyle, Lee Roosevelt, Ashley Lacombe-Duncan, and Nazanin Andalibi. LGBTQ Persons' Pregnancy Loss Disclosures to Known Ties on Social Media: Disclosure Decisions and Ideal Disclosure Environments. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. doi: 10.1145/3411764.3445331. URL https://doi.org/10.1145/3411764.3445331. 2.4.2 (b)

[329] Katyanna Quach. McDonald's AI drive-thru bot accused of breaking biometrics privacy law. *The Register*, June 2021. URL https://www.theregister.com/2021/06/10/mcdonalds_ai_lawsuit/. Accessed June 15, 2023. 2.9.5

[330] Katyanna Quach. AI recruitment software is "automated pseudoscience". *The Register*, October 2022. URL https://www.theregister.com/2022/10/13/ai_recruitment_software_diversity/. 3.3

[331] Lee Rainie, Monica Anderson, Colleen McClain, Emily A. Vogels, and Risa Gelles-Watnick. *AI in Hiring and Evaluating Workers: What Americans Think*. Pew Research Center, April 2023. URL https://www.pewresearch.org/internet/2023/04/20/ai-in-hiring-and-evaluating-workers-what-americans-think/. 3.3, 3.4.1

[332] Fateme Rajabiyazdi, Charles Perin, Jo Vermeulen, Haley MacLeod, Diane Gromala, and Sheelagh Carpendale. Differences that matter: in-clinic communication challenges. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*, page 251–260, Barcelona Spain, May 2017. ACM. ISBN 9781450363631. doi: 10.1145/3154862.3154885. URL https://dl.acm.org/doi/10.1145/3154862.3154885. 5.2

[333] P. A. S. Ralston, J. H. Graham, and J. L. Hieb. Cyber security risk assessment for SCADA and DCS networks. *ISA Transactions*, 46(4):583–594, October 2007. doi: 10.1016/j.isatra.2007.04.003. URL https://www.sciencedirect.com/science/article/pii/S0019057807000754. 4.4.2

[334] Philipp A. Rauschnabel, Alexander Rossmann, and M. Claudia tom Dieck. An adoption framework for mobile augmented reality games: The case of pokémon go. *Computers in Human Behavior*, 76, 2017. doi: 10.1016/j.chb.2017.07.030. 2.4.2

[335] Deepti Balaji Raykar and V Sridhar. Elicitation of personal data categories for implementing data protection: An exploratory study in an educational institution. In *15th Innovations in Software Engineering Conference*, ISEC 2022, 2022. ISBN 9781450396189. doi: 10.1145/3511430.3511443. URL https://doi.org/10.1145/3511430.3511443. 2.4.4

[336] Jeffrey Reaser, Eric Wilbanks, Karissa Wojcik, and Walt Wolfram. *Language Variety in the New South: Contemporary Perspectives on Change and Variation*. UNC Press Books, March 2018. ISBN 9781469638812. Google-Books-ID: l29RDwAAQBAJ. 3.4.4, 3.5.1

[337] RP Reece and Bernd Carsten Stahl. The professionalisation of information security: Perspectives of UK practitioners. *Computers & Security*, 48:182–195, February 2015. doi: 10.1016/j.cose.2014.10.007. 4.4.1

[338] Paul Reilly, Elisa Serafinelli, Rebecca Stevenson, Laura Petersen, and Laure Fallou. Enhancing Critical Infrastructure Resilience Through Information-Sharing: Recommendations for European Critical Infrastructure Operators. In *Transforming Digital Worlds*, Lecture Notes in Computer Science, pages 120–125. Springer International Publishing, March 2018. ISBN 9783319781051. doi: 10.1007/978-3-319-78105-1_15. 4.4.3

[339] Lena Reinfelder, Robert Landwirth, and Zinaida Benenson. Security Managers Are Not The Enemy Either. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–7. Association for Computing Machinery, 2019. ISBN 9781450359702. doi: 10.1145/3290605.3300663. URL https://doi.org/10.1145/3290605.3300663. 4.4.1, 4.4.3

[340] Thomas Reitmaier, Electra Wallington, Dani Kalarikalayil Raju, Ondrej Klejch, Jennifer Pearson, Matt Jones, Peter Bell, and Simon Robinson. Opportunities and challenges of automatic speech recognition systems for low-resource language speakers. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. doi: 10.1145/3491102.3517639. URL https://doi.org/10.1145/3491102.3517639. 3.4.3, 5.3.6

[341] Thomas Reitmaier, Electra Wallington, Ondřej Klejch, Nina Markl, Léa-Marie Lam-Yee-Mui, Jennifer Pearson, Matt Jones, Peter Bell, and Simon Robinson. Situating automatic speech recognition development within communities of under-heard language speakers. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. doi: 10.1145/3544548.3581385. URL https://doi.org/10.1145/3544548.3581385. 3.4.3, 5.3.6

[342] Patrice Renaud, Joanne L Rouleau, Luc Granger, Ian Barsetti, and Stéphane Bouchard. Measuring sexual preferences in virtual reality: A pilot study. *CyberPsychology & Behavior*, 5(1):1–9, 2002. doi: 10.1089/109493102753685836. 2.4.2

[343] Justin Rende. Council post: Why overcoming the cybersecurity labor shortage matters to company success. *Forbes*, March 2023. https://www.forbes.com/sites/forbestechcouncil/2023/03/01/why-overcoming-the-cybersecurity-labor-shortage-matters-to-company-success/. 4.3

[344] Reuters. Bank of Canada says cyber attack could threaten overall financial stability. *Reuters*, May 2023. https://www.reuters.com/article/idUSBCLIGEJ5H/. 4.3

[345] Harvard Law Review. Resetting antidiscrimination law in the age of ai. *Harvard Law Review*, 138(6):1562–1584, April 2025. URL https://harvardlawreview.org/print/vol-138/resetting-antidiscrimination-law-in-the-age-of-ai/. 3.7.4, 6.2.3

[346] Alene Rhea, Kelsey Markey, Lauren D'Arinzo, Hilke Schellmann, Mona Sloane, Paul Squires, and Julia Stoyanovich. Resume format, linkedin urls and other unexpected influences on AI personality prediction in hiring: Results of an audit. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '22, page 572–587, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392471. doi: 10.1145/3514094.3534189. URL https://doi.org/10.1145/3514094.3534189. 3.3, 3.4.1, 3.7.4

[347] Alene K. Rhea, Kelsey Markey, Lauren D'Arinzo, Hilke Schellmann, Mona Sloane, Paul Squires, Falaah Arif Khan, and Julia Stoyanovich. An external stability audit framework to test the validity of personality prediction in AI hiring. *Data Mining and Knowledge Discovery*, 36(6):2153–2193, November 2022. doi: 10.1007/s10618-022-00861-0. URL https://doi.org/10.1007/s10618-022-00861-0. 3.3, 3.4.1, 3.7.4

[348] Tatiana Rice, Kier Lamont, and Jordan Francis. The colorado artificial intelligence act fpf u.s. legislation policy brief, July 2024. URL https://leg.colorado.gov/sites/default/files/images/fpf_legislation_policy_brief_the_colorado_ai_act_final.pdf. 3.7.4, 6.2.3

[349] John Russell Rickford. Unequal partnership: Sociolinguistics and the African American speech community. *Language in Society*, 26(2):161–197, 1997. 3.4.4

[350] David Rind, Sean Lyngaas. Apparent cyberattack forces Florida hospital system to divert some emergency patients to other facilities | CNN Politics. *CNN*, Feb 2023. https://www.cnn.com/2023/02/03/politics/cyberattack-hospital-tallahassee-memorial-florida/index.html. 4.3

[351] Jan Ole Rixen, Mark Colley, Ali Askari, Jan Gugenheimer, and Enrico Rukzio. Consent in the age of ar: Investigating the comfort with displaying personal information in augmented reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022. doi: 10.1145/3491102.3502140. 2.4.2, 2.4.2 (a), 2.4.3, 2.9.4

[352] Paul F. Roberts. Under Scrutiny, Big Ag Scrambles To Address Cyber Risk. *Forbes*, June 2021. https://www.forbes.com/sites/paulfroberts/2021/06/20/under-scrutiny-big-ag-scrambles-to-address-cyber-risk/. 4.3

[353] Adi Robertson. Nreal light review: Hardware is only half the battle. *The Verge*, Nov 2021. URL https://www.theverge.com/22791981/nreal-light-augmented-mixed-reality-glasses-review. Accessed November 30, 2022. 2.4.1 (a)

[354] Jorge A. Rodriguez, Elaine C. Khoong, Stuart R. Lipsitz, Courtney R. Lyles, David W. Bates, and Lipika Samal. Telehealth experience among patients with limited english proficiency. *JAMA Network Open*, 7(5):

e2410691, May 2024. doi: 10.1001/jamanetworkopen.2024.10691. URL https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2818496. 5.3.4

[355] Victoria A. Rodriguez, Elizabeth F. Boggs, Michael C. Verre, Mary Katherine Siebenaler, Jennifer S. Wicks, Cynthia Castiglioni, Hannah Palac, and Craig F. Garfield. Hospital discharge instructions: Characteristics, accessibility, and national guideline adherence. *Hospital Pediatrics*, 12(11):959–970, November 2022. doi: 10.1542/hpeds.2021-006493. URL https://publications.aap.org/hospitalpediatrics/article/12/11/959/189670/Hospital-Discharge-Instructions-Characteristics. 5.3.3

[356] Kat Roemmich, Florian Schaub, and Nazanin Andalibi. Emotion AI at work: Implications for workplace surveillance, emotional labor, and emotional privacy. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, page 1–20. Association for Computing Machinery, April 2023. ISBN 9781450394215. doi: 10.1145/3544548.3580950. URL https://dl.acm.org/doi/10.1145/3544548.3580950. 3.7.4

[357] Franziska Roesner, Tadayoshi Kohno, and David Molnar. Security and privacy for augmented reality systems. *Communications of the ACM*, 57(4), 2014. doi: 10.1145/2580723.2580730. 2.4.2

[358] Danny Ross. Processing biometric data? be careful, under the gdpr. *IAPP*, October 2017. URL https://iapp.org/news/a/processing-biometric-data-be-careful-under-the-gdpr/. 3.7.4

[359] Michael J. Ryan, William Held, and Diyi Yang. Unintended impacts of llm alignment on global representation, 2024. URL https://arxiv.org/abs/2402.15018. 3.4.3

[360] Eric E Sabelman and Roger Lam. The real-life dangers of augmented reality. *IEEE Spectrum*, 52(7):48–53, 2015. doi: 10.1109/MSPEC.2015.7131695. 2.4.1 (a)

[361] Rahime Belen Saglam, Jason R. C. Nurse, and Duncan Hodges. Privacy concerns in chatbot interactions: When to trust and when to worry. In Constantine Stephanidis, Margherita Antona, and Stavroula Ntoa, editors, *HCI International 2021 - Posters*, page 391–399, Cham, 2021. Springer International Publishing. ISBN 9783030786427. doi: 10.1007/978-3-030-78642-7_53. 5.3.8

[362] Nithya Sambasivan, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, and Vinodkumar Prabhakaran. Reimagining algorithmic fairness in india and beyond. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, page 315–328, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383097. doi: 10.1145/3442188.3445896. URL https://doi.org/10.1145/3442188.3445896. 3.7.1

[363] D Samuel. Warren & louis d. brandeis, the right to privacy. *Harvard Law Review*, 4(5):193, 1890. 2.4.4

[364] Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi, and Noah A. Smith. The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, page 1668–1678, Florence, Italy, 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1163. URL https://www.aclweb.org/anthology/P19-1163. 3.4.4

[365] U. Sarkar, A. Karter, Jennifer Y. Liu, N. Adler, Robert Nguyen, Andrea López, and D. Schillinger. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. *J. Am. Medical Informatics Assoc.*, 2011. doi: 10.1136/jamia.2010.006015. 5.3.3

[366] Praveen Kumar Sattarapu, Deepti Wadera, Nguyen Phong Nguyen, Jaspreet Kaur, Sumeet Kaur, and Emmanuel Mogaji. Tomeito or tomahto: Exploring consumer's accent and their engagement with artificially intelligent interactive voice assistants. *Journal of Consumer Behaviour*, 23(2):278–298, 2024. doi: 10.1002/cb.2195. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cb.2195. 3.3, 3.4.3, 5.3.6

[367] Jocelyn Scheirer, Raul Fernandez, and Rosalind W. Picard. Expression glasses: a wearable device for facial expression recognition. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, 1999. doi: 10.1145/632716.632878. 2.4.1 (a)

[368] Yael Schenker, Eliseo J. Pérez-Stable, Dana Nickleach, and Leah S. Karliner. Patterns of interpreter use for hospitalized patients with limited english proficiency. *Journal of General Internal Medicine*, 26(7):712–717, 2011. doi: 10.1007/s11606-010-1619-z. URL http://link.springer.com/10.1007/s11606-010-1

619-z. 5.3.3

[369] Morgan Klaus Scheuerman, Kandrea Wade, Caitlin Lustig, and Jed R Brubaker. How we've taught algorithms to see identity: Constructing race and gender in image databases for facial analysis. *Proceedings of the ACM on Human-computer Interaction*, 4(CSCW1):1–35, 2020. 3.4.2

[370] Arathi Sethumadhavan, Joe Garvin, and Ben Noah. Speech is human and multifaceted: Our approach to studying it should be the same. *Interactions*, 29(5):6–8, August 2022. ISSN 1072-5520. doi: 10.1145/3555049. URL https://doi.org/10.1145/3555049. 3.4.4

[371] Sakib Shahriar, Sonal Allana, Seyed Mehdi Hazratifard, and Rozita Dara. A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle. *IEEE Access*, 11:61829–61854, 2023. doi: 10.1109/ACCESS.2023.3287195. 5.3.8

[372] Ax Sharma. $5.9 million ransomware attack on farming co-op may cause food shortage. *Ars Technica*, September 2021. https://arstechnica.com/information-technology/2021/09/5-9-million-ransomware-attack-on-farming-co-op-may-cause-food-shortage/. 4.3

[373] Renee Shelby, Shalaleh Rismani, Kathryn Henne, AJung Moon, Negar Rostamzadeh, Paul Nicholas, N'Mah Yilla-Akbari, Jess Gallegos, Andrew Smart, Emilio Garcia, and Gurleen Virk. Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23, page 723–741, New York, NY, USA, August 2023. Association for Computing Machinery. ISBN 9798400702310. doi: 10.1145/3600211.3604673. URL https://dl.acm.org/doi/10.1145/3600211.3604673. 3.4.2, 3.7.1

[374] Cong Shi, Xiangyu Xu, Tianfang Zhang, Payton Walker, Yi Wu, Jian Liu, Nitesh Saxena, Yingying Chen, and Jiadi Yu. Face-mic: Inferring live speech and speaker identity via subtle facial dynamics captured by ar/vr motion sensors. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, 2021. doi: 10.1145/3447993.3483272. 2.4.2, 2.4.2 (b)

[375] Jan Simson, Alessandro Fabris, and Christoph Kern. Lazy data practices harm fairness research. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, page 642–659, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704505. doi: 10.1145/3630106.3658931. URL https://doi.org/10.1145/3630106.3658931. 3.4.2

[376] Makoni Sinfree and Pennycook Alastair. *Disinventing and Reconstituting Languages*. Number Vol. 62 in Bilingual Education and Bilingualism. Multilingual Matters, 2006. ISBN 9781853599248. URL https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,sso&db=e000xna&AN=181393&site=ehost-live&scope=site&custid=s8368349. 3.4.4

[377] Mona Sloane. Automation and recruiting: Understanding the intersection of algorithmic systems and professional discretion in the sourcing of job candidates. *SSRN Electronic Journal*, 2023. doi: 10.2139/ssrn.4516453. URL https://www.ssrn.com/abstract=4516453. 3.7.3

[378] Mona Sloane. Boolean Clashes: Discretionary Decision Making in AI-Driven Recruiting. *Communications of the ACM*, April 2025. doi: 10.1145/3708596. URL https://cacm.acm.org/opinion/boolean-clashes-discretionary-decision-making-in-ai-driven-recruiting/. 3.7.3

[379] Mona Sloane, Emanuel Moss, and Rumman Chowdhury. A Silicon Valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns*, 3(2):100425, 2022. doi: 10.1016/j.patter.2021.100425. URL https://linkinghub.elsevier.com/retrieve/pii/S2666389921003081. 3.3, 3.4.1, 3.7.3, 3.7.4

[380] Theodoros Solomou, Ionut-Cristian Canciu, Marios Christodoulou, Panayiotis Savva, Constantinos Yiasemi, Zinonas Antoniou, Ioannis Constantinou, Christos N. Schizas, and Constantinos S. Pattichis. Empowering citizens through translated patient summary access and sharing in digital health ecosystems. *Studies in Health Technology and Informatics*, 316:497–501, August 2024. doi: 10.3233/SHTI240457. 5.3.7

[381] Theodoros Solomou, Christos N. Schizas, and Constantinos S. Pattichis. *Emerging Mobile Health Systems and Services*. IntechOpen, November 2024. ISBN 9780850149449. doi: 10.5772/intechopen.1007846. URL https://www.intechopen.com/chapters/1200746. 5.3.7

[382] Theodoros Solomou, Stelios Mappouras, Efthyvoulos Kyriacou, Ioannis Constantinou, Zinonas Antoniou, Ionut Cristian Canciu, Marios Neophytou, Zoltan Lantos, Christos N. Schizas, and Constantinos S. Pattichis. Bridging language barriers in healthcare: a patient-centric mobile app for multilingual health record access and sharing. *Frontiers in Digital Health*, 7:1542485, February 2025. doi: 10.3389/fdgth.2025.1542485. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11876183/. 5.3.7

[383] Daniel J. Solove. A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154(3):477–564, January 2006. doi: 10.2307/40041279. URL https://www.jstor.org/stable/10.2307/40041279?origin=crossref. 2.4.4, 2.9.1, 5.3.8

[384] Daniel J. Solove. Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, 126:1880, 2012. 2.4.4

[385] Nissy Sombatruang, Tristan Caulfield, Ingolf Becker, Akira Fujita, Takahiro Kasama, Koji Nakao, and Daisuke Inoue. Internet Service Providers' and Individuals' Attitudes, Barriers, and Incentives to Secure IoT. In *32nd USENIX Security Symposium (USENIX Security 23)*, pages 1541–1558. USENIX Association, August 2023. ISBN 978-1-939133-37-3. URL https://www.usenix.org/conference/usenixsecurity23/presentation/sombatruang. 4.4.3

[386] Dipankar Srirag, Nihar Ranjan Sahoo, and Aditya Joshi. Evaluating dialect robustness of language models via conversation understanding, 2024. URL https://arxiv.org/abs/2405.05688. 3.4.3

[387] Dipankar Srirag, Aditya Joshi, and Jacob Eisenstein. Predicting the target word of game-playing conversations using a low-rank dialect adapter for decoder models, 2025. URL https://arxiv.org/abs/2409.00358. 3.4.3

[388] Tim Starks. Analysis | A presidential critical infrastructure protection order is getting a badly needed update, officials say. *Washington Post*, May 2023. https://www.washingtonpost.com/politics/2023/05/11/presidential-critical-infrastructure-protection-order-is-getting-badly-needed-update-officials-say/. 4.3

[389] General Assembly State of Colorado. Senate bill 24-205, May 2024. URL https://leg.colorado.gov/sites/default/files/2024a_205_signed.pdf. 3.7.4, 6.2.3

[390] Alexander Staves, Antonios Gouglidis, and David Hutchison. An Analysis of Adversary-Centric Security Testing within Information and Operational Technology Environments. *Digital Threats: Research and Practice*, 4(1):14:1–14:29, Mar 2023. doi: 10.1145/3569958. URL https://dl.acm.org/doi/10.1145/3569958. 4.4.2

[391] Scott Stein. Varjo's lidar-enabled XR-3 VR headset shows where VR and AR are bound to blend. *CNET*, December 2020. URL https://www.cnet.com/tech/computing/varjos-lidar-enabled-xr-3-vr-headset-shows-where-vr-and-ar-are-bound-to-blend/. Accessed November 30, 2022. 2.4.1 (a)

[392] Scott Stein. Mind control comes to vr, letting me explode alien heads with a thought. *CNET*, Jan 2021. URL https://www.cnet.com/tech/computing/controlling-vr-with-my-mind-nextminds-dev-kit-shows-me-a-strange-new-world/. Accessed November 30, 2022. 2.4.1 (a)

[393] Anne Steketee, Monnica T Williams, Beatriz T Valencia, Destiny Printz, and Lisa M Hooper. Racial and language microaggressions in the school ecology. *Perspectives on Psychological Science*, 16(5):1075–1098, 2021. 3.4.1

[394] Leah Stodart. The best fitness trackers for keeping up with your goals. *Mashable*, Aug 2022. URL https://mashable.com/roundup/wearable-fitness-trackers-guide. Accessed November 30, 2022. 2.4.1 (a)

[395] Nataly R. Espinoza Suarez, Meritxell Urtecho, Samira Jubran, Mei-Ean Yeow, Michael E. Wilson, Kasey R. Boehmer, and Amelia K. Barwise. The roles of medical interpreters in intensive care unit communication: A qualitative study. *Patient Education and Counseling*, 104(5):1100–1108, May 2021. doi: 10.1016/j.pec.2020.10.018. URL https://www.sciencedirect.com/science/article/pii/S073839912030553X. 5.3.3

[396] Hung-Yue Suen, Kuo-En Hung, and Chien-Liang Lin. Intelligent video interview agent used to predict communication skill and perceived personality traits. *Human-centric Computing and Information Sciences*, 10 (1):3, January 2020. doi: 10.1186/s13673-020-0208-3. URL https://doi.org/10.1186/s13673-020-0208-3. 3.3, 3.4.1, 3.7.1

[397] Timothy Summers, Kalle J. Lyytinen, Tony Lingham, and Eugene A. Pierce. How Hackers Think: A Study of Cybersecurity Experts and Their Mental Models. *The Third International Conference on Engaged Management Scholarship*, September 2013. doi: 10.2139/ssrn.2326634. URL https://papers.ssrn.com/abstract=2326634. 4.4.3

[398] Jiao Sun, Thibault Sellam, Elizabeth Clark, Tu Vu, Timothy Dozat, Dan Garrette, Aditya Siddhant, Jacob Eisenstein, and Sebastian Gehrmann. Dialect-robust evaluation of generated text, 2022. URL https://arxiv.org/abs/2211.00922. 3.4.3

[399] Philipp Sykownik, Divine Maloney, Guo Freeman, and Maic Masuch. Something personal from the metaverse: Goals, topics, and contextual factors of self-disclosure in commercial social vr. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022. doi: 10.1145/3491102.3502008. 2.4.2 (b)

[400] Breena R Taira, Vanessa Kreger, Aristides Orue, and Lisa C Diamond. A pragmatic assessment of google translate for emergency department instructions. *Journal of General Internal Medicine*, 36(11):3361–3365, 2021. 5.3.5

[401] Nishtha H. Tandel, Harshadkumar B. Prajapati, and Vipul K. Dabhi. Voice recognition and voice comparison using machine learning techniques: A survey. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2020. doi: 10.1109/ICACCS48705.2020.9074184. 2.4.1 (a)

[402] Wai Yen Tang and Jesse Fox. Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behavior*, 42(6), 2016. doi: 10.1002/ab.21646. 2.4.2 (b)

[403] Matilde Tassinari, Matthias Burkard Aulbach, and Inga Jasinskaja-Lahti. The use of virtual reality in studying prejudice and its reduction: A systematic review. *PLOS ONE*, 17(7), 2022. doi: 10.1371/journal.pone.0270748. 2.4.2 (b)

[404] Rachael Tatman. Gender and dialect bias in YouTube's automatic captions. In Dirk Hovy, Shannon Spruit, Margaret Mitchell, Emily M. Bender, Michael Strube, and Hanna Wallach, editors, *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pages 53–59, Valencia, Spain, April 2017. Association for Computational Linguistics. doi: 10.18653/v1/W17-1606. URL https://aclanthology.org/W17-1606/. 3.4.3, 5.3.6

[405] Jordan Taylor, Wesley Hanwen Deng, Kenneth Holstein, Sarah Fox, and Haiyi Zhu. Carefully unmaking the "marginalized user:" a diffractive analysis of a gay online community. *ACM Trans. Comput.-Hum. Interact.*, June 2024. doi: 10.1145/3673229. URL https://doi.org/10.1145/3673229. Just Accepted. 3.7.2

[406] Gillian Tett. The financial system is alarmingly vulnerable to cyber attack. *Financial Times*, February 2023. https://www.ft.com/content/03507666-aad7-4dc3-a836-658750b880ce. 4.3

[407] Hibby Thach, Samuel Mayworm, Daniel Delmonaco, and Oliver Haimson. (in)visible moderation: A digital ethnography of marginalized users and content moderation on twitch and reddit. *New Media & Society*, Jul 2022. doi: 10.1177/14614448221109804. URL http://journals.sagepub.com/doi/10.1177/14614448221109804. 2.4.2 (b)

[408] Mary Theofanos, Brian Stanton, Susanne Furman, Sandra Spickard Prettyman, and Simson Garfinkel. Be Prepared: How US Government Experts Think About Cybersecurity. In *Workshop on Usable Security (USEC)*. Internet Society, February 2017. URL http://dx.doi.org/10.14722/usec.2017.23006. 4.4.3

[409] Marianthi Theoharidou, Panayiotis Kotzanikolaou, and Dimitris Gritzalis. Risk assessment methodology for interdependent critical infrastructures. *International Journal of Risk Assessment and Management (IJRAM)*, 15(2/3):128, 2011. doi: 10.1504/IJRAM.2011.042113. URL http://www.inderscience.com/link.php?id=42113. 4.4.2

[410] Frédérique Thonon, Swati Perrot, Abhijna Vithal Yergolkar, Olivia Rousset-Torrente, James W Griffith,

Olivier Chassany, and Martin Duracinsky. Electronic tools to bridge the language gap in health care for people who have migrated: Systematic review. *Journal of Medical Internet Research*, 23(5):e25131, May 2021. doi: 10.2196/25131. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8138704/. 5.2

[411] Lisa Collins Tidwell and Joseph B. Walther. Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research*, 28(3):317–348, 2002. doi: 10.1111/j.1468-2958.2002.tb00811.x. URL https://academic.oup.com/hcr/article/28/3/317-348/4330958. 5.3.8

[412] Atnafu Lambebo Tonja, Fazlourrahman Balouchzahi, Sabur Butt, Olga Kolesnikova, Hector Ceballos, Alexander Gelbukh, and Thamar Solorio. Nlp progress in indigenous latin american languages, May 2024. URL http://arxiv.org/abs/2404.05365. arXiv:2404.05365 [cs]. 3.7.2

[413] Yuan-Chi Tseng, Weerachaya Jarupreechachan, and Tuan-He Lee. Understanding the benefits and design of chatbots to meet the healthcare needs of migrant workers. *Proc. ACM Hum.-Comput. Interact.*, 7(CSCW2), October 2023. doi: 10.1145/3610106. URL https://doi.org/10.1145/3610106. 5.3.7

[414] U.S. Department of Commerce U.S. Census Bureau. Place of birth for the foreign-born population in the united states. U.S. Census Bureau, . URL https://data.census.gov/table/ACSDT5Y2023.B05006?q=American+Community+Survey+5-year+data&t=Foreign-Born&y=2023&d=ACS+5-Year+Estimates+Detailed+Tables. Accessed on 25 May 2025. 5.2, 5.3.1

[415] U.S. Department of Commerce U.S. Census Bureau. Place of birth by year of entry for the foreign-born population. U.S. Census Bureau, . URL https://data.census.gov/table/ACSDT5Y2023.B05015?q=American+Community+Survey+5-year+data&t=Foreign-Born&y=2023&d=ACS+5-Year+Estimates+Detailed+Tables. Accessed on 25 May 2025. 5.2, 5.3.1

[416] Susana Valdez and Ana Guerberof-Arenas. "google translate is our best friend here": A vignette-based interview study on machine translation use for health communication. *Translation Spaces*, April 2025. doi: 10.1075/ts.24040.val. URL http://www.jbe-platform.com/content/journals/10.1075/ts.24040.val. 5.3.5

[417] Mina Valizadeh, Xing Qian, Pardis Ranjbar-Noiey, Cornelia Caragea, and Natalie Parde. What clued the AI doctor in? on the influence of data source and quality for transformer-based medical self-disclosure detection. In Andreas Vlachos and Isabelle Augenstein, editors, *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 1201–1216, Dubrovnik, Croatia, May 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.eacl-main.86. URL https://aclanthology.org/2023.eacl-main.86/. 5.3.8

[418] Varjo. Varjo and OpenBCI Partner to Bring Neurotechnology to Spatial Computing, May 2022. URL https://varjo.com/company-news/openbci-and-varjo-partner-to-bring-neurotechnology-to-spatial-computing/. Accessed November 30, 2022. 2.4.1 (a)

[419] Varjo. Terms of Service for Varjo XR-3 and VR-3, 2022. URL https://varjo.com/terms-of-service-for-varjo-xr-3-and-vr-3/. Accessed November 30, 2022. 2.4.1 (b)

[420] Varjo. Eye tracking, 2022. URL https://varjo.com/use-center/get-to-know-your-headset/eye-tracking/. Accessed November 30, 2022. 2.4.1 (a)

[421] Varjo. Privacy policy, 2022. URL https://varjo.com/privacy-policy/. Accessed November 30, 2022. 2.4.1 (b)

[422] Lucas Nunes Vieira, Minako O'Hagan, and Carol O'Sullivan. Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases. *Information, Communication & Society*, 24(11):1515–1532, August 2021. doi: 10.1080/1369118X.2020.1776370. URL https://www.tandfonline.com/doi/full/10.1080/1369118X.2020.1776370. 5.3.5

[423] Daniel Votipka, Rock Stevens, Elissa Redmiles, Jeremy Hu, and Michelle Mazurek. Hackers vs. Testers: A Comparison of Software Vulnerability Discovery Processes. In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 374–391. IEEE Computer Society, 2018. 4.4.3

[424] Daniel Votipka, Kelsey R. Fulton, James Parker, Matthew Hou, Michelle L. Mazurek, and Michael Hicks.

Understanding security mistakes developers make: Qualitative analysis from Build It, Break It, Fix It. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 109–126. USENIX Association, August 2020. ISBN 978-1-939133-17-5. URL https://www.usenix.org/conference/usenixsecurity20/presentation/votipka-understanding. 4.4.3

[425] Esma Wali, Yan Chen, Christopher Mahoney, Thomas Middleton, Marzieh Babaeianjelodar, Mariama Njie, and Jeanna Neefe Matthews. Is machine learning speaking my language? a critical look at the nlp-pipeline across 8 human languages, July 2020. URL http://arxiv.org/abs/2007.05872. arXiv:2007.05872 [cs]. 3.4.4

[426] Robert Walton. A month after "malicious" cyberattack, a small colorado utility still doesn't have all systems back online. *Utility Dive*, December 2021. https://www.utilitydive.com/news/a-month-after-malicious-cyberattack-a-small-colorado-utility-still-doesn/610983/. 4.3

[427] Michael Walzer. *Spheres of justice: A defense of pluralism and equality*. Basic Books, 1983. ISBN 0-465-08189-4. 2.9.2

[428] Jihong Wang. 'I only interpret the content and ask practical questions when necessary.' Interpreters' perceptions of their explicit coordination and personal pronoun choice in telephone interpreting. *Perspectives*, 29 (4):625–642, July 2021. doi: 10.1080/0907676X.2018.1549087. URL https://www.tandfonline.com/doi/full/10.1080/0907676X.2018.1549087. 5.3.3

[429] Azmine Toushik Wasi, Taki Hasan Rafi, and Dong-Kyu Chae. Diaframe: A framework for understanding bengali dialects in human-ai collaborative creative writing spaces. In *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing*, CSCW Companion '24, page 268–274, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400711145. doi: 10.1145/3678884.3681862. URL https://doi.org/10.1145/3678884.3681862. 3.7.2

[430] Jess Weatherbed. Texas is replacing thousands of human exam graders with ai, April 2024. URL https://www.theverge.com/2024/4/10/24126206/texas-staar-exam-graders-ai-automated-scoring-engine. 3.3

[431] Suzanne Weisband and Sara Kiesler. Self disclosure on computer forms: meta-analysis and implications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '96, page 3–10, New York, NY, USA, 1996. Association for Computing Machinery. ISBN 0897917774. doi: 10.1145/238386.238387. URL https://doi.org/10.1145/238386.238387. 5.3.8

[432] Frederike Wenzlaff, Peer Briken, and Arne Dekker. Video-based eye tracking in sex research: A systematic literature review. *The Journal of Sex Research*, 53(8), 2016. doi: 10.1080/00224499.2015.1107524. 2.4.2

[433] Nicole Wetsman. Cyberattack delays patient care at major US hospital chain. *The Verge*, October 2022. https://www.theverge.com/2022/10/11/23398707/cyberattack-hospital-system-patient-care-issues. 4.3

[434] Whiteford, Lisa Bauner, and Slaven Jesic. Client alert: Avoiding legal pitfalls and risks in workplace use of artificial intelligence, September 2024. URL https://www.jdsupra.com/legalnews/client-alert-avoiding-legal-pitfalls-2004765/. 3.7.4, 6.2.3

[435] Cedric Deslandes Whitney and Justin Norman. Real risks of fake data: Synthetic data, diversity-washing and consent circumvention. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '24, page 1733–1744, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704505. doi: 10.1145/3630106.3659002. URL https://doi.org/10.1145/3630106.3659002. 6.3.2

[436] Flynn Wolf, Adam J. Aviv, and Ravi Kuber. Security Obstacles and Motivations for Small Businesses from a CISO's Perspective. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 1199–1216. USENIX Association, August 2021. ISBN 978-1-939133-24-3. URL https://www.usenix.org/conference/usenixsecurity21/presentation/wolf. 4.4.3

[437] Marilyn Wolf and Dimitrios Serpanos. Safety and Security in Cyber-Physical Systems and Internet-of-Things Systems. *Proceedings of the IEEE*, 106(1):9–20, January 2018. doi: 10.1109/JPROC.2017.2781198. 4.4.1

[438] Walt Wolfram. The sociolinguistic construction of African American. *The Oxford handbook of African American Language*, 338, 2015. 3.4.4

[439] Kathryn A Woolard. Language ideology. *The International Encyclopedia of Linguistic Anthropology*, pages 1–21, 2020. doi: 10.1002/9781118786093.iela0217. 3.7.2

[440] Qiushi Wu, Yue Xiao, Xiaojing Liao, and Kangjie Lu. OS-Aware Vulnerability Prioritization via Differential Severity Analysis. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 395–412. USENIX Association, August 2022. URL https://www.usenix.org/conference/usenixsecurity22/presentation/wu-qiushi. 4.4.2

[441] Ziang Xiao, Wesley Hanwen Deng, Michelle S. Lam, Motahhare Eslami, Juho Kim, Mina Lee, and Q. Vera Liao. Human-Centered Evaluation and Auditing of Language Models. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI EA '24, page 1–6, New York, NY, USA, May 2024. Association for Computing Machinery. ISBN 9798400703317. doi: 10.1145/3613905.3636302. URL https://doi.org/10.1145/3613905.3636302. 3.7.2

[442] Jing Yang, Yen-Lin Chen, Lip Yee Por, and Chin Soon Ku. A systematic literature review of information security in chatbots. *Applied Sciences*, 13(11):6355, May 2023. doi: 10.3390/app13116355. URL https://www.mdpi.com/2076-3417/13/11/6355. 5.3.8

[443] Farah Yousry. Cyberattacks on health care are increasing. Inside one hospital's fight to recover. *NPR*, May 2023. https://www.npr.org/sections/health-shots/2023/05/08/1172569347/cyberattacks-on-health-care-are-increasing-inside-one-hospitals-fight-to-recover. 4.3

[444] Aijia Yuan, Edlin Garcia Colato, Bernice Pescosolido, Hyunju Song, and Sagar Samtani. Improving workplace well-being in modern organizations: A review of large language model-based mental health chatbots. *ACM Trans. Manage. Inf. Syst.*, 16(1), February 2025. ISSN 2158-656X. doi: 10.1145/3701041. URL https://doi.org/10.1145/3701041. 5.3.8

[445] Marco Zappatore and Gilda Ruggieri. Adopting machine translation in the healthcare sector: A methodological multi-criteria review. *Computer Speech & Language*, 84:101582, 2024. doi: 10.1016/j.csl.2023.101582. URL https://linkinghub.elsevier.com/retrieve/pii/S0885230823001018. 5.3.5, 5.3.6

[446] Amir H. Zargarzadeh and Anandi V. Law. Access to multilingual prescription labels and verbal translation services in california. *Research in Social and Administrative Pharmacy*, 7(4):338–346, 2011. doi: 10.1016/j.sapharm.2010.08.001. URL https://linkinghub.elsevier.com/retrieve/pii/S155174111000080X. 5.3.3

[447] Rena Zendedel, Barbara C. Schouten, Julia C. M. Van Weert, and Bas Van Den Putte. Informal interpreting in general practice: the migrant patient's voice. *Ethnicity & Health*, 23(2):158–173, February 2018. doi: 10.1080/13557858.2016.1246939. URL https://www.tandfonline.com/doi/full/10.1080/13557858.2016.1246939. 5.3.3

[448] Kim Zetter. Hacking Wall Street. *The New York Times*, July 2021. https://www.nytimes.com/2021/07/03/business/dealbook/hacking-wall-street.html. 4.3

[449] Jia Zhang, Yinian Zhou, Rui Xi, Shuai Li, Junchen Guo, and Yuan He. Ambiear: Mmwave based voice recognition in nlos scenarios. volume 6, 2022. doi: 10.1145/3550320. 2.4.1 (a)

[450] Kexin Zhang, Elmira Deldari, Zhicong Lu, Yaxing Yao, and Yuhang Zhao. "it's just part of me:" understanding avatar diversity and self-presentation of people with disabilities in social virtual reality. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, 2022. doi: 10.1145/3517428.3544829. 2.4.2 (b)

[451] Shikun Zhang, Yan Shvartzshnaider, Yuanyuan Feng, Helen Nissenbaum, and Norman Sadeh. Stop the spread: A contextual integrity perspective on the appropriateness of covid-19 vaccination certificates. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, 2022. doi: 10.1145/3531146.3533222. 2.4.4

[452] Ioannis Zografopoulos, Juan Ospina, Xiaorui Liu, and Charalambos Konstantinou. Cyber-Physical Energy Systems Security: Threat Modeling, Risk Assessment, Resources, Metrics, and Case Studies. *IEEE Access*,

9:29775–29818, 2021. doi: 10.1109/ACCESS.2021.3058403. 4.4.2

[453] Jonathan Zong and J. Nathan Matias. Data refusal from below: A framework for understanding, evaluating, and envisioning refusal as design. *ACM Journal on Responsible Computing*, 1(1):1–23, March 2024. doi: 10.1145/3630107. URL https://dl.acm.org/doi/10.1145/3630107. 3.7.2

[454] Ethan Zuckerman. *117. Alondra Nelson, Biden's Head of Science and Technology Policy, talks AI, Trump's research funding cuts, and how memes replaced Happy Days*. Number 117 in Initiative for Digital Public Infrastructure at UMass Amherst. Initiative for Digital Public Infrastructure at UMass Amherst, July 2025. URL https://publicinfrastructure.org/podcast/117-alondra-nelson/. 6.3.2