

**Elicitation and Empathy with AI-enhanced Adaptive Assistive Technologies
(AATs): Towards Sustainable Inclusive Design Method Education**

*Nora McDonald, Aaron Massey, Foad Hamidi **

ABSTRACT

Efforts to include people with disabilities in design education are difficult to scale, and dynamics of participation need to be carefully planned to avoid putting unnecessary burdens on users. However, given the scale of emerging AI-enhanced technologies and their potential for creating new vulnerabilities for marginalized populations, new methods for generating empathy and self-reflection in technology design students (as the future creators of such technologies) are needed. We report on a study with Information Systems graduate students where they used a participatory elicitation toolkit to reflect on two cases of end-user privacy perspectives towards AI-enhanced tools in the age of surveillance capitalism: their own when using tools to support learning, and those of older adults using AI-enhanced adaptive assistive technologies (AATs) that help with pointing and typing difficulties. In drawing on the experiences of students with intersectional identities, our exploratory study aimed to incorporate intersectional thinking in privacy elicitation and further understand its role in enabling sustainable, inclusive design practice and education. While aware of the risks to their own privacy and the role of identity and power in shaping experiences of bias, students who used the toolkit were more sanguine about risks faced by AAT users—assuming more data equates to better technology. Our tool proved valuable for eliciting reflection but not empathy.

Keywords: Adaptive Assistive Technology, Privacy, Intersectionality, Design Justice, Participatory Toolkit, Elicitation Toolkit, Problem-based Learning, Computing higher education

* Nora McDonald, Department of Information Science and Technology, George Mason University, United States
Email: norakmcdonald@gmail.com
Aaron Massey, Information Systems Department, University of Maryland, Baltimore County, United States
Email: akmassey@umbc.edu
Foad Hamidi, Information Systems Department, University of Maryland, Baltimore County, United States
Email: foadhamidi@umbc.edu

INTRODUCTION

AI-enhanced Adaptive Assistive Technologies (AATs) that collect user data to improve functionality, present potential benefits and usability gains for people with disabilities (Hamidi et al., 2018). These systems can also pose risks to users who are particularly vulnerable to privacy violations and resulting harms (McDonald et al., 2021). However, it is unclear how future creators of such technologies (i.e., students) should be sensitized to the complex design tradeoffs these systems pose. In this paper, we explore this space in the context of problem-based learning in higher education and make several contributions. First, we describe a novel, and in-progress, approach (toolkit plus intersectional method) to eliciting intersectional privacy considerations about AI-enhanced AATs and other AI-enhanced technologies with students. Second, we respond to the growing need to incorporate non-normative methods into privacy design thinking for populations with disabilities, particularly when those designs are based on software already in use (where vulnerabilities may be less visible or taken for granted). Third, we reflect on the shortcomings of the activities' toolkit for empathic use without, as student participants suggest, direct contact with vulnerable AAT users, and offer recommendations for improving our process. Fourth, we present a problem-based learning solution that balances the increasing call for more inclusive and participant-driven intersectional design with the realities of classroom (and corporate) settings. Ultimately, our goal is to develop a toolkit and an accompanying methodology to capture the privacy-related tensions around disability and AI so that designers can better account for the needs of users with disabilities using a sustainable method. Our approach is meant to complement, rather than replace, existing face-to-face design and educational methods by investigating new approaches that allow for nuanced, ongoing, and participatory dialogue between technology designers and user populations. We aim to support empathy and perspective-taking while still being mindful of the burden that participation may place on users.

The potential for AI-enhanced systems to compound intersectional discriminations for those with disabilities or with changing health conditions is increasingly being recognized (Whittaker et al., 2019). For instance, data collected by AATs could be used by third-parties in ways that may limit opportunities or lead to other harms for these individuals. We are motivated to investigate what happens when a system could determine that you had visual impairments *and* were a minority simply by using your typing data and triangulating it with other data. What *then* if that information got into the hands of a bad actor, or an employer, or an insurance company? The discriminatory possibilities compound, yet the extent to which vulnerable users grasp or feel capable of mitigating these risks remains unclear.

Consideration of the possible harms caused by AI-enabled systems raises questions about how to teach a new generation of developers about designing for those with complex abilities in a way that is, both sensitive to their unique needs and intersectional vulnerabilities and respectful of their investment (of time, emotion, etc.). It also offers valuable opportunities to engage students in problem-based learning projects that encourage reflection and consideration of real-world scenarios. Researchers like Shoshana Costanza-Chock (Costanza-Chock, 2020) have articulated important agendas for inclusive design and the need for human-computer interaction (HCI) scholars to find ways to integrate those ideas into scalable and sustainable paradigms for learning and designing. Moreover, the benefits of problem-based learning approaches in computing pedagogy (Karan & Brown, 2022), particularly in integrating critical thinking and consideration of societal problems (Scholkmann et al., 2023), have been demonstrated.

Indeed, in the study of older adults with disabilities, the emphasis on technology innovation has traditionally been on inclusivity, but the growth of AI applications requires a broader focus on other harms that could affect these intersectional populations. Where before, the concern may have been focused on accessibility, it is now also about being the target of discrimination by insurance companies and advertisers. There is growing urgency for sustainable and scalable interventions in the classroom to introduce ethics curriculum to future designers of these systems. However, there are some challenges:

First, the current state of AI ethics education is underdeveloped and suffers from a lack of attention and coverage in academia (Saltz et al., 2019). This has repercussions for the technologies that are developed in the market, where scholars point to limited thinking in the way that AI might behave after deployment and the people it may harm (Webb, 2019). While computer science educators have long acknowledged the importance of ethics, only recently has there been a demand for a rigorous and integrated curriculum—and even that presents challenges for instructors who may not be equipped to incorporate ethics in their classes (Fiesler et al., 2020). Perhaps as a result, when ethical perspectives are taught, they are often standalone courses and not embedded throughout the curriculum (Fiesler et al., 2020).

Second, widespread, indiscriminate data collection poses serious ethical considerations beyond simply enabling an AI system (Mak, 2018). Storing, managing, and securing data is not a trivial undertaking, making it a rich but challenging area to draw on for problem-based learning projects that are rooted in current and real-world issues.

Third, the systems that help those in need of AATs are increasingly built on existing data and capabilities that, when enhanced, might breed yet more paradoxical relationships to our technology: the more adaptive and helpful, the greater the potential for harm. The “dual use” nature (e.g., technology that connects us can be used to surveil us) of our technologies presents important opportunities and risks (Chatterjee et al., 2018).

Relatedly, technologies reify and reproduce inequalities and heteropatriarchal norms in part because of the data they are trained on and which is reinforced time and again by study (e.g., through A/B testing) of their existing user base. Yet, while there is support for personalized systems that break with the model of the single user, personalization has its downsides—e.g., surveillance, tracking, and filter bubbles (Costanza-Chock, 2020).

Building on our previous use of this elicitation toolkit with AAT users (Hamidi et al., 2020; McDonald et al., 2021), this study looks specifically at international and/or minority technology students—both of whom experience risk with adaptive text AI because of their marginalized identities. Among issues that overlap for both older AAT users and non-native language speakers is the way in which AI-generated language tools could both normalize speech and reduce credibility when their users are very dependent on them (Hancock et al., 2020). Both of these communities interact with technologies that collect and adapt to personal data with a kind of reliance that falls outside what is normative because of different challenges (e.g., vision and mobility vs. language and cultural pressures). We consider that despite their distinct experiences and identities, students and AAT users share in common relationships to surveillance power and risk, through algorithmic technology, relationships that innovative design learning methodologies could draw on.

With this study, we investigated the possibilities and limitations of a participatory approach for enabling technology students to consider and empathize with the perspectives of a vulnerable population—older adults who use AATs. We found that students expect that a system designed for AAT users would likely collect the same amount of data, or possibly more for older AAT users. And while older AAT users told us in our previous research (Hamidi et al., 2020) that they assume that an adaptive AI system would easily identify their disability or disease, students did not. Even those who did entertain that assumption felt that it would benefit, not harm, the user. In contrast to what we learned in our previous research with older adults with Essential Tremors (ET) (Hamidi et al., 2020), students were more certain about institutional surveillance but believed that there was nothing they or anyone else could do about it. Students all agreed the activities toolkit was helpful in eliciting reflections about themselves. However, the students also reflected on its limits in allowing them to consider risk on others' behalf without engaging with them personally.

In the following sections, we first report on the AAT privacy research landscape and describe the intersectionality conceptual frame that we utilized alongside this toolkit to elicit thinking on AAT users' behalf. We then describe an elicitation toolkit and accompanying interview instrument we have developed for students' intersectional reflections. We next present results from our interview study using the toolkit with seven IS graduate students. We conclude with implications for future iterations of the toolkit.

RELATED LITERATURE

AATs and Data Capitalism/Colonialism

The possible scope of discrimination resulting from the widespread deployment of AI/ML systems is, indeed, alarming. The AI/ML technologies used in policing and immigration enforcement (Adzima, 2017; Buolamwini, 2019; Ferguson, 2017; Hao, 2018), crime risk assessment (Angwin, 2015; Richardson, 2014), and welfare benefits (Eubanks, 2006, 2018) exacerbate inequalities like income, gender identity, race, and class. Most critical to our understanding of the tensions in the new AI horizon is that they cannot be solved merely with technical approaches (*Algorithmic Accountability Policy Toolkit*, 2018; Hagendorff, 2020; Wong, 2020). That is because these tensions emerge from a dominant social and political matrix that necessitates awareness of social and political context and an understanding of power (*Algorithmic Accountability Policy Toolkit*, 2018) and data/surveillance capitalism or colonialism—the mechanism of capitalism and the colonizing impulses that undergird the treatment of user data as commodities to extract, trade, exploit, and sell (Couldry & Mejias, 2019; West, 2019; Zuboff, 2015, 2019). When user data are repackaged, they profile groups of individuals based on socio-economics, race/ethnicity, and other identity vulnerabilities. According to a US Senate Report, a data broker creates and sells consumer groups based on, for example, financial vulnerability, ethnicity, and age with categories like: “Rural and Barely Making It” and “Ethnic Second-City Strugglers” (*A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes*, 2013).

End-user Perspectives toward Adaptive and Personalized Technologies

A growing body of research on user perceptions and attitudes towards privacy tradeoffs of adaptive and personalized applications has emerged in the last few decades (e.g., Ur et al., 2012). Many of these projects focus on online marketing (Shklovski et al., 2014), Internet-of-Things (IoT) and wearable applications for health (Gorm & Shklovski, 2016; Zhou & Piramuthu, 2014). In several studies, users expressed feelings of “creepiness” when they learned (or considered) how their data could be used outside of the original context of an application’s use (Angulo & Ortlieb, 2015; Seberger et al., 2022; Shklovski et al., 2014; Ur et al., 2012). Researchers have identified a mismatch between users’ mental models and how personal data is actually collected that can lead to unpleasant surprises and discomfort when users learn about discrepancies between their expectations and the actual privacy characteristics of an application (Gorm & Shklovski, 2016; Kang et al., 2015; Ur et al., 2012; Zhou & Piramuthu, 2014), though scholars argue, convincingly, that this disconnect is more a product of manufactured “resignation” than misunderstanding (Draper & Turow, 2019; Seberger et al., 2021; Seberger et al., 2022). While awareness of bias with commercially available AI among disabled users was found to be low, the same users express discomfort with AI systems’ collection of personal data

and its use by third parties and institutions and are worried about the discovery of hidden disabilities (Park et al., 2021).

AI-enhanced Technologies and Ethics in Education and Design

As emerging technologies increasingly integrate AI and other automation systems in design, scholars have begun to warn of the lack of ethical education in the technology design and development curriculum (Hagendorff, 2020; McDonald et al., 2022; Webb, 2019). Also frequently missing from conversations about AI ethics in classrooms is a consideration for technologies that impose discriminatory practices once AI is deployed (Whittaker et al., 2018), and the way in which political systems and attendant moral norms and deliberations shape AI. One of these norms is the idea that “free” AI-enhanced tools (email, grammar tools, etc.) come with a price—user information like contacts, content, and web activity. What is missing from much of AI ethics is an understanding of the relationships of power in which AI systems are situated and the contexts in which they interact with individuals (Hagendorff, 2020; McDonald et al., 2022).

When AI designers adapt software engineering best practices, they fail to appreciate the difficulty of ethically managing vulnerable populations with intersectional concerns. For example, the use of explicit personas, rather than roles, to better understand stakeholder concerns has been used widely in industry for almost two decades (Miller & Williams, 2006). Use of personas has proven so successful that software engineers now apply it to refining their own development processes (Ford et al., 2017). However, after years of use in industry, when researchers examined how personas applied to stakeholders with complex disability identities, the technique was found to be inadequate (Edwards et al., 2020).

According to Costanza-Chock, “far too often user personas are created out of thin air by members of the design team (if not autogenerated by services like Userforge), based on their own assumptions or stereotypes about groups of people who might occupy a very different location in the matrix of domination”—i.e., their relationship to power is different and likely privileged (Costanza-Chock, 2020). Costanza-Chock is equally critical of “disability simulation,” where a non-disabled person navigates a space as if they were disabled to locate challenges and elicit empathy. First, they cause researchers and designers to respond to their own experience, subverting the experience of those that designers intend to help. If something seems real, we may be even more inclined to “turn it off,” to distance ourselves from the disability that we have the privilege to remove with, say, our blindfold. Alternatively, we might overstate certain constraints while overlooking others. Bennet and Rosner and Edwards et al. say that while empathy activities and personas are important, there is a difference between “being like” (helping) vs “being with” (supporting and empowering) (Bennett & Rosner, 2019; Edwards et al., 2020). Although well-intentioned, empathy exercises and personas reduce disability to obvious and ergonomic constraints, and distance us from or subvert the disabled other,

they may fail to take into account the overlapping oppression of identity and disability in relation to structural inequality.

Notably, a number of approaches have been developed for assistive technology and accessibility design education, that recognize the importance of including people with disabilities at every stage of the design process. These approaches include User-Sensitive Inclusive Design (Newell et al., 2011), Design for User Empowerment (Ladner, 2015), and Ability-based Design (Wobbrock et al., 2011). Shinohara et al. developed an approach, Design for Social Accessibility, that recognizes the importance of supporting student awareness of socially usable aspects of a design in addition to its functionality (Shinohara et al., 2018). Additionally, recognizing the importance of empathy as a key aspect in accessibility design, this approach calls for the inclusion of perspectives from users with and without disabilities in the design process, and the use of methods that support consideration of social factors in accessible design (Shinohara et al., 2018).

While efforts to include people with disabilities in design education are effective, they are difficult to scale, and also dynamics of participation by people with disabilities need to be carefully planned to avoid putting unnecessary burdens on users solely for the benefit of students. There is, indeed, no substitute for the “real” user with disabilities, but their experiences might get abstracted as design activities advance and form the basis of innovative problem-based learning experiences connected to real-world issues. These considerations motivated us to see how we could explore complementary ways to insert user perspectives in the process and contribute to more sustainable, inclusive design methods, while also better understanding the limitations that these approaches may entail. Deciding who is *most* responsible (designers, institutions, or regulators) is beyond the scope of this paper. However, we do focus on influencing the ethical process of students who will someday be designers employed by technology companies designing AI.

Given this landscape, there is a need to study how to elicit intersectional thinking from vulnerable, current, and future technologists on behalf of vulnerable others who may use their technologies, with the goal of moving closer to developing methods that sensitize students and AI design practitioners to structural inequalities. In the next section, we discuss how intersectionality builds on this inclusive perspective to argue that we must seek out the perspectives of those who are not privileged and understand their experiences of power.

Conceptual Lens: Intersectionality

Intersectionality can play an important role in helping students understand the way AI technologies exacerbate discrimination—particularly for those who are disabled with other identity vulnerabilities—through profiling and surveillance by powerful institutions (Collins, 2019; Collins & Bilge, 2016; Crenshaw, 1989, 1991; Eubanks, 2018). It offers a useful framework for designing the current study to explore how AI students approach

the relationship between their experience of identities and institutions that impose power and those of other vulnerable populations.

Collins positions intersectionality as a theory that is perpetually becoming—and a “way of thinking” (Collins, 2019). In that spirit, we adopt Collin’s matrix of domination, a paradigm that focuses on how power is organized, and which integrates with her thinking about intersectionality. We use it to understand whether experiences with a intersectional oppression can sensitize student designers to non-normative identities. The core constructs of Collins’ matrix of domination are: interpersonal (how people’s actions shape power relations), disciplinary (which rules apply, to whom, and when; e.g., bureaucratic organizations perform routine surveillance for the sake of efficiency), hegemonic or cultural (conditions under which power takes hold) and structural (how powerful institutions are organized; e.g., laws, policies, etc.) domains of power (Collins, 1990, 2019; Collins & Bilge, 2016). Domains of power (particularly disciplinary and structural domains) usefully describe the way we engaged (or had hoped) to engage users with our elicitation tool to find common ground around power mechanisms that may act on different identities. Disciplinary domains and structural domains are shaped by business logics and privacy-invasive algorithms that fuel the accumulation of data for advertising. One way that the matrix manifests is through surveillance capitalism, reifying inequality through algorithmic-enabled surveillance that disproportionately harms certain marginalized groups.

A PARTICIPATORY ACTIVITIES TOOLKIT FOR ELICITING END-USER PRIVACY PERSPECTIVES AND INTERSECTIONAL REFLECTIONS

To facilitate conversations about privacy considerations for diverse and vulnerable individuals, we adapted a toolkit used in (Hamidi et al., 2020) consisting of a set of low-fidelity cards, strips, and charts. In this study, we rendered the toolkit pictured in Figure 1 as a remote tool (the only viable way to interview students during COVID-19), using software elements designed in Axure Share (*Axure Share.*, n.d.) to help students think about privacy considerations on behalf of themselves and on behalf of vulnerable AAT user groups—those with mobility and vision impairments. We chose to characterize AAT users as experiencing difficulties with typing and pointing devices due to mobility or vision impairments, though we refer to them as AAT users throughout this paper. Our motivation was to deemphasize having a specific condition to help to focus more on the experience of having difficulty when accessing a computer (i.e., pointing, typing, seeing). Figure 1 (Right) shows the paper version of the toolkit we adapted for this study (please see (Hamidi et al., 2020) for details of how the toolkit works). We also created a remote version of this toolkit (Figure 1, Left) that formed the basis of the one we used in the current study with the students.

Adaptive Assistive Technology (AAT) Prototypes

The remote kit (adapted from Figure 1, Left) includes two software prototypes that represent AATs that participants might use to help access the web. The first system is the popular cloud-based writing assistant, Grammarly (*Grammarly.*, n.d.). For our second system, we used the Pointing Interaction Notifications and AdapTations (PINATA) (Hamidi et al., 2018) to help users who experience difficulty when using pointing devices. It consists of a dynamic bubble cursor (Grossman & Balakrishnan, 2005) that simulates the functionality of dynamically changing size in response to users’ pointing performance and the location of the cursor. PINATA monitors a user’s pointing behavior over time, and when errors are detected (e.g., a link is missed while it is being clicked) increases the size of the cursor.

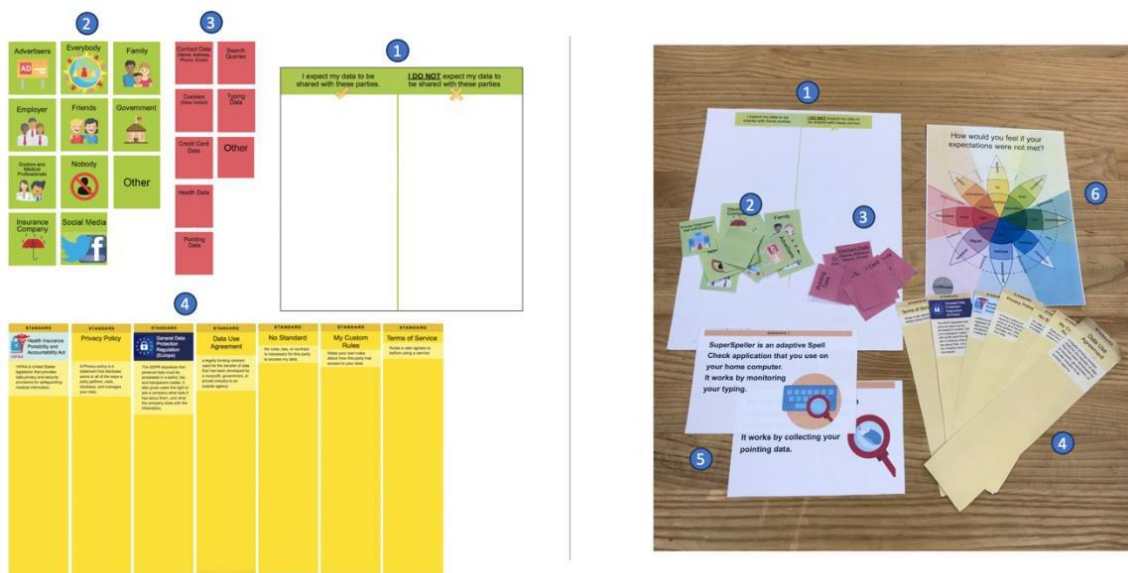


Figure 1. Two versions of the Participatory Activities Toolkit (paper version, on the right; remote version on the left). The paper version was developed for our earlier study with ET older adults. The toolkit included the following elements: (1) Expectations Chart, (2) Third Party Cards, (3) Data Type Cards, (4) Privacy Standard Strips, (5) Scenario Cards (replaced in our study with application demos), (6) Wheel of Emotion (adapted in our study with students as a verbal exercise to elicit intersectional reflections about disability and powerful institutions that linked to students’ own experience of discrimination in the context of data capitalism).

Participatory Activity Procedures

The kit includes a series of four activities (described below) for three scenarios:

- (1) a Grammarly system built for students,
- (2) a Grammarly system for older individuals with vision/mobility impairment that use AATs, and
- (3) PINATA system for older individuals with vision/mobility impairment or Essential Tremors (ET) that use AATs.

We described ET to students as a condition that can make it difficult to steady one's hands and control their cursor.

Activity 1: What data should an AAT collect?

We first asked the participant what types of data they expected the AAT we demo'ed to collect. We gave them the **red Data Type Cards** and asked them whether they expect the application to collect them or not. We asked participants to elaborate on why each of the data types would be collected by the application in the given scenario.

Activity 2: Who should access my/their data?

We next gave them **green Third-Party Cards** and asked participants to place them on the **Expectations Chart** to indicate which parties they expected had access to their data collected by the application in the given scenario. We asked them to explain their reasoning when placing the green Third-Party Cards in the chart. For both activities, blank cards were offered if participants wanted to include new items.

Activity 3: What standard(s) should protect my/their data?

In the last activity, participants selected **yellow Privacy Standard Strips** to protect their data and were asked to explain which standards they would like enforced.

Activity 4: Eliciting Intersectional Reflections.

To elicit the intersectional reflections of technology students, we asked them to consider the scenarios of (1) the Grammarly system being designed for them, (2) Grammarly being designed for an older individual experiencing typing or pointing difficulties, and (3) PINATA being designed for an older individual experiencing typing or pointing difficulties. We had students reflect on their own privacy vulnerabilities (particularly, those stemming from their visa status). We then asked them to "imagine" being visually or motor impaired and what that might mean for dependency on technology and also on the privacy risks. We asked them to reflect on how AAT users would "feel" about data collection and access by third-parties given their own experiences of surveillance with them. Our goal was to facilitate students drawing parallels with the intersectional perspectives of vulnerable individuals with respect to how automated and adaptive systems may collect both their data and who might have access to that data in ways that could result in harms.

Participants and Interview Procedures

We recruited 7 university students (5 international from Asian countries; all of non-white ethnicity; all under 30; 3 females) who were completing a degree in Information Systems at our university and were taking a course in algorithm design at the time of the study. While demographic information can be relevant in qualitative research, we have decided to include only summary data since they were attending a small class, and triangulating their data may result in de-anonymization. We conducted remote, semi-structured video

interviews with students that lasted an average of 59 minutes. We refer to these participants as G1-G7 in our reporting of results.

Data Collection and Analysis

We audio recorded and transcribed each session using a video conferencing system and took screenshots of participants' completed Expectation Charts. Notes and memos were taken before and after each interview. We took an iterative thematic analysis approach to identify and synthesize themes within the interview transcriptions. For each study, the team member who conducted the interviews reviewed transcripts and notes and wrote memos which were organized into themes that were both emergent and based on the interview framing. The interviewers revisited videos and transcripts of key noted themes and anecdotes for use in this paper. All research was approved by our institutional regulation board (IRB).

RESULTS

Already our Privacy is Exposed to Everybody

Overall, technology student participants assume the same amount of data collected for them by Grammarly and PINATA would be collected for AAT users. When it comes to considerations about who the data is being shared with, students expressed concerns about being targeted by the government and advertisers, but these concerns were reserved largely for themselves and not AAT users. Our analysis shows how students refer to various encounters with power in relation to their privacy but do not imagine that those risks or harms exist for AAT users. Students did not worry about government access to AAT users' data and assumed that targeted advertising and oversight would be welcome to these groups. When given the scenario for PINATA, students expressed even fewer concerns about privacy for AAT users—rather, they just saw the software as only helpful.

Grammarly

We asked students to consider both the case of older adults using Grammarly and themselves using the application. When considering their own use scenario, students expected that Grammarly collects their typing data (the content of what they write), but they don't connect these mechanisms with monitoring and profiling of AAT users. Most also assumed it collected their contact data, and potentially their cookies and search history. In their view, the data collected about them by Grammarly was comparable to, or maybe just a little less than, what is collected for AAT users, whom they tend to believe, either would not be harmed or not understand enough to worry or care.

If students did worry about the nefarious use of data collected by Grammarly, it was in the context of their use. For instance, one student worried about their data being used for visa and immigration surveillance, which has caused them to limit or eliminate its use in

certain settings: “There are spaces that are more intimate or more privacy sensitive, and one of those is email.” [G3]. These students generally understood themselves to be operating in an unequal playing field, where their activities could be more heavily scrutinized and used against them in ways, unlike their privileged counterparts. By contrast, even if students suspected that AAT users might have some reservations about the collection of their data, their expectation was that they would “make peace” [G6] with the data being used because they rely on it. These misconceptions about people with disabilities needing AATs, even at the cost of privacy, are reminiscent of misconceptions previously identified in accessibility research (Shinohara & Wobbrock, 2011).

Some students found reassurance in their belief that a system like Grammarly could not detect that a user has a disability or a medical condition, though they did imagine that it could detect language ability, including being foreign or young. One student spontaneously pointed out that Grammarly could not tell the difference between, say, an AAT user who was blind versus a non-native English speaker, or someone with learning disabilities. They reasoned that because so many identities are possibly mistaken for vision impairment, having a separate user interface might be necessary to learn more about the user’s health condition to improve the system and also for advertisers to more effectively advertise vision- related items: “It could be a child. It could be a person from another country whose first language is not English ... If you have the second version of your software that is solely for the people with vision impairment, then, you know the degree of his impairment ... and based on that, you're sending the advertisement to that particular person” [G3]. Only one student believed that Grammarly could reveal patterns of behavior through pointing data (which they assumed would be collected for Grammarly) that would expose disability, leading to opportunity loss and other discriminations: “If they're tracking the pointing data of users, they know the patterns of their users, and it can be used as evidence of impairment ... If this is exposed to other organizations, those companies could deny them health insurance, position of employment, that kind of thing” [G5].

PINATA

We did not ask participants to think about how PINATA would use *their* data, only how it would use the data of someone who experiences pointing difficulties. Students typically assumed that PINATA would collect more data than Grammarly. For instance, they tended to agree that PINATA would have to collect cookies in order to work because the system would have to know what links were clicked. One student speculated that if PINATA worked by image processing, then it would collect all the red Data Cards provided in the toolkit.

While a few concluded that PINATA would collect clicking and/or image data, only one pointed out that it would make them very uncomfortable if the tool did this (“Taking

images from my computer. I'm very uncomfortable" [G5]). The other students seemed to entertain the possibility that all sites collect those data for marketing purposes. The student who did express discomfort went on to say that they are helpless to do anything about who sees their data: "I still feel uncomfortable, but honestly ... our privacy is exposed to everybody" [G5].

For the most part, students' discomfort reflected their own privacy concerns and not necessarily those of AAT users. When thinking about users with pointing difficulties, students expected PINATA data to be shared with more parties—e.g., healthcare professionals, family, and friends—than would be shared by Grammarly. They expect that doctors, family, and friends would want to know if the individual was progressing. One student expected the government would also want to know how many people have the condition saying: "[The] government needs to know that" [G4].

With PINATA, concern for the medical needs of individuals overrode any concerns about privacy. It may be that students viewed PINATA as akin to medical equipment and, thus, subject to a set of standards different from those used for technologies they use, available to the general public.

Regulations

We asked students to talk about what regulations they wanted to impose on these tools. Discussions about regulations did not prompt increased concern for individuals with disabilities. If anything, some students were convinced that tools designed specifically for disabilities—like PINATA—are harmless and, thus, would consider imposing fewer regulations: "If nothing bad happens, then a data use agreement is not needed either." [G6]

Domains of Power

Students convey implicit awareness of *disciplinary domains of power* (what rules apply, to whom, and when) as well as *structural domains of power* (how immigration institutions operate and use AI infrastructures). They nevertheless do not consider these domains of power for older AAT users. For instance, they don't consider the monitoring of text, typing, and pointing data by insurance companies and other advertisers profiling to offer different services as problematic. For instance, one student reflected on how these data are useful for national security: "The most obvious one is national security ... If we're looking at AI, Grammarly, and stuff like that, any sequence of words that might raise a red flag in terms of national security, that could be information that the government cares about" [G1].

Knowledge of *structural domains of power* did cause students to alter their practices, but they did not imagine that these same structures (e.g., in the form of insurance companies or advertisers) threatened AAT users.

Positive, not Intersectional Thinking

Students identify power imbalance in their own interaction with AI and powerful institutions that deploy it but not necessarily in those of AAT users; and even when they do, they assume that these privacy breaches are all for the best. For instance, students assume that personalization is not only an appropriate privacy tradeoff for someone with disabilities but may not consider it a tradeoff at all. The following student expressed a sentiment shared by others that while they wouldn't want to share certain data with PINATA, for a user with disabilities, more data would improve the tool: "You ended up feeling like, 'no, maybe I don't want to share that information for myself.' But if I had visual impairment, I would want to share more information to make the system work" [G4].

Students might have been moved to think about the struggle of individuals who rely on adaptive systems, but they tended to suppose that data would ameliorate these problems. For example, they imagined that the data these tools collected would improve the experience and provide the benefit of interventions by doctors, and families, as well as more targeted advertising. By way of rationalization, students also seemed convinced (paradoxically) that the data these tools collect would not expose *specific* vulnerabilities (despite having just speculated about the benefits such a system would impart by inferring this information) and thus did not represent a risk. Only one student worried that the data these systems collected could be used for discriminatory practices, like denying health insurance. While a few students considered that an adaptive system might be able to determine that one was having "difficulties," they were fairly confident that it would *not* be able to tell the difference between, say, visual impairment or being a non-native speaker, or having dyslexia.

There is a Limit to "Being With" Other's Oppressions using this Tool

Student participants found the elicitation tool helpful in making them reflect deeply about data use and regulations: "I did enjoy the interview because it definitely opened my mind. So just kind of like ... aimlessly using technology, especially if it's backed by AI" [G1]. They considered the activities toolkit effective at making them think differently about how they designed AI systems, but also, admittedly, limited in helping them struggle with taking other's perspective—with being in the other persons' shoes: "I cannot step forward [in their shoes] by considering all these issues" [G4].

Yet some asserted that if they were to design a product for people with visual or motor disabilities, they would want to have them on the team (or interview them) to understand their "struggles" and "stories" as well as their "needs": "For other people? No, I feel like I don't know their struggles or their stories. You would need to have someone like that on board ... If you actually have someone who is visually impaired or semi visually impaired, at least then you know, and you'll have a better understanding of what their needs

are” [G2]. One posited that they would do this in an iterative way to understand their comfort with the system at each step: “I would have to ask them their expectations, then design. After designing, I would go to them and I ask them if they are comfortable, then we can go to the final system” [G4].

Students’ obliviousness to the potential misuses of technology for AAT users, coupled with a genuine and deep desire to involve the vulnerable AAT end-user in design, has several implications for this toolkit going forward, which we will discuss in the next section.

DISCUSSION AND IMPLICATIONS FOR ITERATION

While students often imagined that Grammarly and PINATA collected the same data for different populations (e.g., both themselves and AAT users), they also tended towards thinking that the more disabled the user, the more important and helpful it would be to collect and utilize their data. They described, for instance, how having access to more data would help advertisers target people with disabilities with products that were better geared toward them and would help the AI work better. Others considered that more data would be useful to doctors or governments who wanted to monitor the condition of users with disabilities and possibly intervene. This notion of technology as being only of service, rather than also potentially harmful, to people with disabilities, might be the product of an idealistic frame of mind that a game-like elicitation activity potentiated, or it might be a kind of rationalization in the face of domains and structures of power they feel helpless to control.

RELATIONSHIP TO INTERSECTIONAL NOTIONS OF STRUCTURES OF POWER AND DATA/SURVEILLANCE CAPITALISM/COLONIALISM

Most of our student participants were adamant that third-party organizations collect data. Students sometimes, when prompted, noted their discomfort with the amount of data that may be collected about them—how it flows (*structural domain of power*) and what might be done with it (*disciplinary domain of power*)—but largely considered that there is nothing they can do about it. Students’ ideas about powerful institutional oversight evidence certain contradictions. They expressed concern about government and advertiser oversight *of them* (though not of AAT users) while, at the same time, expressing resignation about that oversight, deeming it a given, something that cannot be helped or stopped and even perhaps a legitimate “right,” whatever uncomfortable effects it might have on their behaviors. As they shift the focus of conversation from themselves to AAT, however, they seem to conclude that the use is simultaneously less capable of individuation, and also more benign. Very few considered the ways in which systems of power might abuse the data collected about vulnerable users. For them, the *cultural*

domain of power rested on the idea that those with disabilities operate in a world where the data collected about them is not intended for nefarious use. This framing was “plausible” because, in their view, these technologies were designed to help AAT users, while they could be used to jeopardize their stay in the US.

Future research might explore educating students about surveillance capitalism so that they can more readily make the connection between the use of surveillance and data by the government (they experience) and its use by insurance companies and advertisers (which AAT users experience)—perhaps even also drawing connections with this disciplinary power to structures in our laws, policy, and economy that systematically overlook those who stand to lose the most. Students who did think about the potential for these systems to identify users’ disabilities, were convinced it would be indistinguishable from other vulnerabilities and, therefore, not a problem. Putting the best case on it, students imagined these tools would collect data that led to improvements for AAT users. Perhaps including interactive elements in the toolkit that combine data types and illustrate how they may be triangulated by powerful institutions or stakeholders can help students reflect on the implications for intersectional identities.

INTERSECTIONAL FRAMING AND FUTURE TOOLKIT DESIGN AND RESEARCH

We attempted to connect the experience of algorithmic oppression by one group to their interpretation of another’s and found these experiences were not readily transferrable. Students ultimately wanted direct contact with those whom they were designing for.

What is not clear is whether having overlapping experiences of oppression can be leveraged. One step towards that end would be to *more clearly delineate and elucidate mechanisms and risks in our design and activities*. In future iterations of this tool and protocol, we need to do a better job of eliciting consideration of identities and power. Our intention with the tool is not only to enlighten student technologists about the experiences of vulnerable populations through problem-based learning, rather, it is also to create opportunities for reflection on one’s own experiences and how they may relate to those of others who experience privacy threats in relation to AI-enabled systems. We hope this will lead to increased empathy.

LIMITATIONS AND REFLECTIONS

We have just discussed several ways to “do better” in future studies. But this research also raises some uncomfortable questions. For one, should we be looking to marginalized individuals to recognize the struggle of others simply because they experience similar forms of violation (e.g., surveillance of their typing data that has more salient or meaningful consequences)? Second, information Systems and HCI graduate programs may also historically be partly to blame for instilling students with a notion that all

technology means well. It is thus no surprise that the students we spoke with expected that designs for accessibility were made with the best intentions. Do we need to radically reframe our mission and/or curriculum to account for the ineluctable harms caused by technology in the context of our current economies and political systems? Finally, the onus cannot be on those who suffer disproportionately to be the agents of change (and be the subject of research that singles them out) and so studies like this must ultimately find ways to negotiate parallels with those who are more privileged. Social justice in design must be a multi-pronged effort.

CONCLUSIONS

We explored the utility of a privacy elicitation toolkit with graduate information systems students and found it useful to elicit reflections about the risks of collecting data to enable AI technologies on a student's own behalf but not necessarily on behalf of others—real-world AAT users. Students were sometimes able to associate their own risks with aspects of their identity that leave them vulnerable but did not extrapolate those identities or vulnerabilities to diverse real-world AAT users. Future work should explore interview, and activity prompts that focus more on identity-based exploration, for instance, incorporating more intersectional interview methods where one's identity and contexts of power are linked to specific experiences of risk (Windsong, 2018) as well as tools for eliciting empathy and thinking about risk, like scenarios. We heard from several student participants that being able to actually talk to individuals they were trying to imagine would help to improve the tool. We will need to explore ways of incorporating feedback from AAT communities into our elicitation activities while still adhering to our goal of sustainable methods.

References

- A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes.* (2013). [Staff Report for Chairman Rockefeller]. United States Senate Committee on Commerce, Science, and Transportation.
- Adzima, C. (2017, June 15). *Using Amazon Rekognition to Identify Persons of Interest for Law Enforcement.* Amazon Web Services.
<https://aws.amazon.com/blogs/machine-learning/using-amazon-rekognition-to-identify-persons-of-interest-for-law-enforcement/>
- Algorithmic Accountability Policy Toolkit.* (2018). AI Now Institute.

- Angulo, J., & Ortlieb, M. (2015). “WTH..!?!” *Experiences, Reactions, and Expectations Related to Online Privacy Panic Situations*. 19–38.
<https://www.usenix.org/conference/soups2015/proceedings/presentation/angulo>
- Angwin, J. (2015, September 10). *First Library to Support Tor Anonymous Internet Browsing Effort Stops After DHS Email*. ProPublica.
<https://www.propublica.org/article/library-support-anonymous-internet-browsing-effort-stops-after-dhs-email>
- Axure Share. (n.d.). Retrieved May 3, 2020, from <https://share.axure.com/>
- Bennett, C. L., & Rosner, D. K. (2019). The Promise of Empathy: Design, Disability, and Knowing the “Other.” *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300528>
- Buolamwini, J. (2019, April 24). *Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces*. Medium.
<https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced>
- Chatterjee, R., Doerfler, P., Orgad, H., Havron, S., Palmer, J., Freed, D., Levy, K., Dell, N., McCoy, D., & Ristenpart, T. (2018). The Spyware Used in Intimate Partner Violence. *2018 IEEE Symposium on Security and Privacy (SP)*, 441–458.
<https://doi.org/10.1109/SP.2018.00061>
- Collins, P. H. (1990). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (1st edition). Routledge.
- Collins, P. H. (2019). *Intersectionality as Critical Social Theory*. Duke University Press Books. Collins, P. H., & Bilge, S. (2016). *Intersectionality* (1 edition). Polity.
- Costanza-Chock, S. (2020). *Design Justice: Community-Led Practices to Build the Worlds We Need*. The MIT Press.
- Couldry, N., & Mejias, U. A. (2019). Data Colonialism: Rethinking Big Data’s Relation to the Contemporary Subject. *Television & New Media*, 20(4), 336–349.
<https://doi.org/10.1177/1527476418796632>
- Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine. *University of Chicago Legal Forum*, 1989(1).
- Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, 43(6), 1241–1299.
- Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society*, 21(8), 1824–1839. <https://doi.org/10.1177/1461444819833331>
- Edwards, E. J., Monet, C. M., & Branham, S. M. (2020). Three Tensions Between Personas and Complex Disability Identities. *Conference on Human Factors in Computing Systems Extended Abstracts*. CHI’20, Honolulu, USA.

- Eubanks, V. (2006). Technologies of Citizenship: Surveillance and Political Learning in the Welfare System. In T. Monahan (Ed.), *Surveillance and Security*. Routledge. <https://doi.org/10.4324/9780203957257-12>
- Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- Ferguson, A. G. (2017). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. NYU Press.
- Fiesler, C., Garrett, N., & Beard, N. (2020). *What Do We Teach When We Teach Tech Ethics?: A Syllabi Analysis*. 289–295. <https://doi.org/10.1145/3328778.3366825>
- Ford, D., Zimmermann, T., Bird, C., & Nagappan, N. (2017). Characterizing software engineering work with personas based on knowledge worker actions. *Proceedings of the 11th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, 394–403. <https://doi.org/10.1109/ESEM.2017.54>
- Gorm, N., & Shklovski, I. (2016). Sharing Steps in the Workplace: Changing Privacy Concerns Over Time. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4315–4319. <https://doi.org/10.1145/2858036.2858352>
- Grammarly. (n.d.). Retrieved May 3, 2020, from <https://www.grammarly.com/>
- Grossman, T., & Balakrishnan, R. (2005). *The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area*. 281–290. <https://doi.org/10.1145/1054972.1055012>
- Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*. <https://doi.org/10.1007/s11023-020-09517-8>
- Hamidi, F., Poneris, K., Massey, A., & Hurst, A. (2018). Who Should Have Access to My Pointing Data?: Privacy Tradeoffs of Adaptive Assistive Technologies. *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 203–216. <https://doi.org/10.1145/3234695.3239331>
- Hamidi, F., Poneris, K., Massey, A., & Hurst, A. (2020, April). Using a Participatory Activities Toolkit to Elicit Privacy Expectations of Adaptive Assistive Technologies. *W4A '20*.
- Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations. *Journal of Computer-Mediated Communication*, 25(1), 89–100. <https://doi.org/10.1093/jcmc/zmz022>
- Hao, K. (2018). Amazon is the invisible backbone behind ICE's immigration crackdown. *MIT Technology Review*. <https://www.technologyreview.com/2018/10/22/139639/amazon-is-the-invisible-backbone-behind-ices-immigration-crackdown/>

- Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). "My Data Just Goes Everywhere." User Mental Models of the Internet and Implications for Privacy and Security. *Eleventh Symposium On Usable Privacy and Security SOUPS'15*, 35–52.
<https://www.usenix.org/conference/soups2015/proceedings/presentation/kang>
- Karan, E., & Brown, L. (2022). Enhancing Students' Problem-solving Skills through Project-based Learning. *Journal of Problem Based Learning in Higher Education*, 10(1), 74-87. <https://doi.org/10.54337/ojs.jpblhe.v10i1.6887>
- Ladner, R. E. (2015). Design for user empowerment. *Interactions*, 22(2), 24–29.
<https://doi.org/10.1145/2723869>
- Mak, A. (2018, April 19). *How a JPMorgan Security Team Reportedly Used Palantir's Tools to Spy on Its Employees*. Slate Magazine.
<https://slate.com/technology/2018/04/jpmorgan-used-palantir-tools-monitor-employee-activity-bloomberg-report.html>
- McDonald, N., Akinsiku, A., Hunter-Cevera, J., Sanchez, M., Kephart, K., Berczynski, M., & Mentis, H. M. (2022). Responsible Computing: A Longitudinal Study of a Peer-led Ethics Learning Framework. *ACM Transactions on Computing Education*, 22(4), 47:1-47:21. <https://doi.org/10.1145/3469130>
- McDonald, N., Massey, A., & Hamidi, F. (2021). AI-Enhanced Adaptive Assistive Technologies: Methods for AI Design Justice. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 4(Special Issue on Responsible AI and Human-AI interaction).
- Miller, R. L., & Williams, L. L. (2006). *Personas: Moving beyond role-based requirements engineering*.
- Newell, A. F., Gregor, P., Morgan, M., Pullin, G., & Macaulay, C. (2011). User-Sensitive Inclusive Design. *Universal Access in the Information Society*, 10(3), 235–243. <https://doi.org/10.1007/s10209-010-0203-y>
- Park, J. S., Bragg, D., Kamar, E., & Morris, M. R. (2021). Designing an Online Infrastructure for Collecting AI Data From People With Disabilities. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 52–63. <https://doi.org/10.1145/3442188.3445870>
- Richardson, J. (2014). Spinoza, Feminism and Privacy: Exploring an Immanent Ethics of Privacy.
Feminist Legal Studies; Dordrecht, 22(3), 225–241.
<http://dx.doi.org.ezproxy2.library.drexel.edu/10.1007/s10691-014-9271-3>
- Saltz, J., Skirpan, M., Fiesler, C., Gorelick, M., Yeh, T., Heckman, R., Dewar, N., & Beard, N. (2019). Integrating Ethics Within Machine Learning Courses. *ACM Trans. Comput. Educ.*, 19(4), 32:1-32:26. <https://doi.org/10.1145/3341164>

- Scholkmann, A., Stegeager, N., & Miller, R. K. (2023). Integrating the Integration: The Role of Problem-Based Learning in Bringing Together Social Science and Humanities (SHH) and Science, Technology, Engineering and Mathematics (STEM). *Journal of Problem Based Learning in Higher Education*, 11(1), 98-123. <https://doi.org/10.54337/ojs.jpblhe.v11i1.7371>
- Seberger, J. S., Llavore, M., Wyant, N. N., Shklovski, I., & Patil, S. (2021). Empowering Resignation: There's an App for That. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445293>
- Seberger, J. S., Shklovski, I., Swiatek, E., & Patil, S. (2022). Still Creepy After All These Years: The Normalization of Affective Discomfort in App Use. *CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3491102.3502112>
- Shinohara, K., Bennett, C. L., Pratt, W., & Wobbrock, J. O. (2018). Tenets for Social Accessibility: Towards Humanizing Disabled People in Design. *ACM Transactions on Accessible Computing*, 11(1), 6:1-6:31. <https://doi.org/10.1145/3178855>
- Shinohara, K., & Wobbrock, J. O. (2011). In the shadow of misperception: Assistive technology use and social interactions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 705–714. <https://doi.org/10.1145/1978942.1979044>
- Shklovski, I., Mainwaring, S. D., Skúladóttir, H. H., & Borgthorsson, H. (2014). Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use. *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, 2347–2356. <https://doi.org/10.1145/2556288.2557421>
- Ur, B., Leon, P. G., Cranor, L. F., Shay, R., & Wang, Y. (2012). Smart, useful, scary, creepy: Perceptions of online behavioral advertising. *Proceedings of the Eighth Symposium on Usable Privacy and Security*, 1–15. <https://doi.org/10.1145/2335356.2335362>
- Webb, A. (2019). *The Big Nine: How the Tech Titans and Their Thinking Machines Could Warp Humanity*. PublicAffairs.
- West, S. M. (2019). Data Capitalism: Redefining the Logics of Surveillance and Privacy. *Business & Society*, 58(1), 20–41. <https://doi.org/10.1177/0007650317718185>
- Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., Meredith, R. M., Rankin, J., Rogers, E., Sala, M., & Sarah, M. W. (2019). *Disability, Bias, and AI*. AI Now Institute.

- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S. M., Richardson, R., Schultz, J., & Schwartz, O. (2018). *AI Now 2018 Report*. AI Now Institute.
- Windsong, E. A. (2018). Incorporating intersectionality into research design: An example using qualitative interviews. *International Journal of Social Research Methodology*, 21(2), 135–147. <https://doi.org/10.1080/13645579.2016.1268361>
- Wobbrock, J. O., Kane, S. K., Gajos, K. Z., Harada, S., & Froehlich, J. (2011). Ability-Based Design: Concept, Principles and Examples. *ACM Transactions on Accessible Computing*, 3(3), 9:1- 9:27. <https://doi.org/10.1145/1952383.1952384>
- Wong, P.-H. (2020). Democratizing Algorithmic Fairness. *Philosophy & Technology*, 33(2), 225–244. <https://doi.org/10.1007/s13347-019-00355-w>
- Zhou, W., & Piramuthu, S. (2014). Security/privacy of wearable fitness tracking IoT devices. *2014 9th Iberian Conference on Information Systems and Technologies (CISTI)*, 1–5. <https://doi.org/10.1109/CISTI.2014.6877073>
- Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, 30(1), 75–89.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (1 edition). PublicAffairs.