CISE HCC CAREER: Towards a Critical AI Literacy for Emotion AI

Noura Howell, PhD, Assistant Professor, Georgia Institute of Technology, nhowell8@gatech.edu

Overview

Emotion AI is an emergent, controversial set of techniques that (claim to) classify human psychological states. Emotion AI is proposed, deployed, and sometimes withdrawn after criticism, in high stakes contexts such as education, hiring, or work. Computer scientists, psychologists, and policy experts debate potential benefits and harms, arguing for polarized futures ranging from panoptic expansion to complete bans. Proponents expound benefits for well-being and security, while critics decry privacy and civil liberties risks, cultural bias, and shaky scientific foundations. Undergirding these polarized debates seem to be competing imagined futures of what Emotion AI can/not and should/not do. Without articulating these envisioned futures for Emotion AI, hasty flawed Emotion AI deployments may cause harm, garner criticism, and ultimately hamper AI innovation. Articulating more diverse, beneficial, and ethical futures for Emotion AI can help Emotion AI deliver on its promise for societal benefit and advance responsible AI innovation.

My long-term goal is to articulate futures from diverse stakeholders, to shape inclusive, beneficial AI innovation. Amidst polarized debate on Emotion AI, I side with neither extreme; I work to surface diverse perspectives that may guide Emotion AI development or advocate for curtailment or bans. The **objective** in this proposal is to identify Emotion AI futures with (Group 1) socioeconomically marginalized teens and young adults, and (Group 2) technology workers in computing or design roles, because (G1) are disproportionately impacted by AI bias while (G2) are uniquely positioned to improve or refuse to build Emotion AI. **Outputs** of the proposed research include (Output I) a taxonomy of desirable and undesirable futures with Emotion AI, (Output II) identified failure modes of Emotion AI, and (Output III) design recommendations for more ethical Emotion AI systems. These outputs will be useful to technology workers in Emotion AI, and educators for AI literacy and AI ethics.

The perspectives of (Group 1) and (Group 2) are essential. (G1) Socioeconomically marginalized teens and young adults navigate biased, opaque AI systems with increasing Emotion AI surveillance in work, education, and hiring, while (G2) technology workers increasingly advocate for ethics in the products they build. I have been studying the tech industry's engagement with Emotion AI for about a decade through design futuring methods that seek to rework dominant techno-narratives and articulate diverse alternative futures. I partner with community groups on STEM outreach. Understanding how (G1) and (G2) imagine Emotion AI futures is key to shaping technology development and curriculum for more beneficial Emotion AI futures.

Methods: We will host a series of speculative design workshops and post-workshop interviews. Speculative design workshops will invite (G1) and (G2) to collectively imagine alternative futures with Emotion AI and create representations of these imagined futures. Interviews will ask (G1) and (G2) to individually articulate their future visions for Emotion AI, ask them to reflect on potential failures and harms of Emotion AI, and invite them to share design recommendations for Emotion AI. We will use critical discourse analysis to analyze the representations to discover how these groups imagine key stakeholders, activities, and benefits/harms of Emotion AI to develop (Output I) futures representations. We will use thematic analysis of interviews to discover how these groups anticipate failures and harms of Emotion AI to produce (OII) failure modes and (OIII) ethical design recommendations.

Intellectual Merit

Related work studies how a general public imagines Emotion AI, shows that Emotion AI demands unique considerations, and traces imagined futures as pivotal in the development of Emotion AI.

This proposal addresses a gap of studying imagined futures of Emotion AI with (G1) socioeconomically marginalized teens and young adults and (G2) technology workers in computer and design roles, with an explicit emphasis on potential failure modes and harms of Emotion AI. Project outputs of (OI) futures representations, (OII) failure modes, and (OIII) ethical design recommendations will be useful for technology development and AI ethics curriculum.

Broader Impacts

Emotion AI is increasingly deployed in everyday life as surveillance that people cannot avoid in work, education, transit, social media, hiring, or other everyday contexts. This project addresses an urgent moral imperative to better understand how people want or do not want to live alongside Emotion AI. The speculative design workshops will offer enhanced AI and STEM literacy for (G1) socioeconomically marginalized teens and young adults. Creating representations in the workshops will involve tinkering with code, offering CS learning for (G2). The workshops will also offer a much-needed space for (G2) tech workers to engage in ethical reflections offsite from their employer and the constraints of their direct professional responsibilities.

CISE HCC CAREER: Towards a Critical AI Literacy for Emotion AI

Project Description

1 Introduction

Emotion AI claims to predict and classify sensitive psychological inferences such as categories of emotion, alertness, or stress, or even evaluate job candidates or predict criminal intent (Crawford et al. 2019; Marda and Ahmed 2021). Emotion AI is controversial. Many applications endanger civil liberties (ibid) and continue to be adopted despite renowned computer scientists describing these predictive knowledge claims in no uncertain terms as 'AI snake oil' that does not and fundamentally cannot work (Narayanan 2021). Yet, other application of Emotion AI have demonstrated benefits for mental health (Slovak et al. 2022; R. Picard 2022; Sas et al. 2020; Sanches, Janson, et al. 2019). With Emotion AI projected to nearly double to USD 43B between 2022-2027 (*Markets* 2022), there is an urgent need to shape the rapid growth of Emotion AI toward beneficial societal impacts. For this, what is needed is a deeper understanding of what kinds of applications, scenarios, and ways of living with Emotion AI people do and do not want. Without this understanding, polarized debates calling for either panoptic expansion or outright bans of Emotion AI are unlikely to enact practical, pragmatic change that preserves benefits for mental health while protecting civil liberties.

The proposed research investigates sociotechnical imaginaries of Emotion AI to understand what ways of living with Emotion AI people do and do not want. **Outputs of the proposed research include a taxonomy of desirable and undesirable scenarios with Emotion AI, and identified failure modes, harm vectors, and strategies for prevention, mitigation, or redressing these failures and harms.** Emotion AI can fail in unique ways, and have unique forms of algorithmic harm, due to the sociocultural complexity of emotion. **These outputs will be useful to** policymakers regulating Emotion AI applications and data protections, engineers and designers deciding what Emotion AI applications to build and how, as well as educators integrating Emotion AI into AI literacy and AI ethics education programs. The proposed work is part of my broader research agenda employing design futuring as a social scientific method to investigate societal and ethical implications of Emotion AI, especially biopolitical contestations around how and by whom Emotion AI knowledge claims are produced and sanctioned. My work surfaces assumptions undergirding Emotion AI and traces how and by whom Emotion AI materials, devices, and techniques are designed, developed, adopted, or rejected. My work attends keenly to marginalization, algorithmic oppression (Noble 2018), and biometric surveillance (Browne 2018).

This work builds on existing STS scholarship on Emotion AI. *DigitalSTS: a field guide for science and technology studies*, argues "the digitally mediated quantification and expression of human emotions is an increasingly central part of the experience of digitally mediated existence. As a discipline, science and technology studies (STS) is ideally placed to explore and critique the ways the varied landscape of human emotional experience and expression shifts as a 'matter of concern' refracted through diverse digital interfaces, systems, and platforms" (Stark 2019, p.117). Emotion AI often aims to classify emotions into discrete categories, and such "classifications should be recognized as the significant site of political and ethical work that they are" (Bowker and Star 1999, p.320). A workshop co-hosted by Stark, Nafus, myself, and others calls for critically imagining roles of biodata and Emotion AI in daily life (Merrill et al. 2017). By investigating sociotechnical imaginaries of Emotion AI, the proposed research aims to uncover ethical and biopolitical aspects of these technologies. Sociotechnical imaginaries are "collectively held and performed visions of desirable futures (or of resistance against the undesirable), and they are also animated by shared

understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology" (Jasanoff and Kim 2015, p.19). Sociotechnical imaginaries can help analyze ethics and power of science and technology, and help differentiate between collective narratives and individual subjectivity (ibid). One of the most effective approaches to analyzing sociotechnical imaginaries is comparison (ibid, pp.24-25).

The proposed research will compare sociotechnical imaginaries of Emotion AI put forth by Emotion AI companies and AI education programs, and compare with the potentially more individualized imaginaries of (1) technology workers in technical, design, or management roles who work on Emotion AI systems, features, techniques, or applications and (2) undergraduate and graduate students in computer science, design, human-computer interaction, and related fields.[1] These groups are interesting to compare because current students face increasing surveillance in education, hiring, and social media compared to technology workers, yet these students are poised to become the next generation of technology workers who can develop or resist future Emotion AI applications.

The proposed research will use **critical discourse analysis to examine institutional sociotechnical imaginaries of companies and education programs, as well as interviews and speculative design workshops to study individual imaginaries of these workers and students**. We will compare institutional sociotechnical imaginaries with the individual voices of workers and students. Critical discourse analysis will examine public-facing materials from (1) Emotion AI companies and (2) AI education programs, to articulate institutionalized imaginaries of Emotion AI. Critical discourse analysis can be effective for surfacing sociotechnical imaginaries of emergent technologies (Wong and Mulligan 2016). Interviews with (1) workers and (2) students will elicit individual imaginaries of Emotion AI, inviting interviewees to describe their hopes and fears for future ways of living with Emotion AI, their perspectives on potential benefits and harms with Emotion AI, and especially how they consider ethical risks of marginalization and ways of mitigating such risks. Speculative design workshops will invite participants to imagine alternative futures with Emotion AI, discuss the ethical implications of these futures, and articulate their values and preferences for future ways of living with Emotion AI. Speculative design is an emergent cross-disciplinary way for STS to engage information technologies (Vertesi et al. 2019).

The **intellectual merit** of this work stems from its unique focus on technology workers' and students' perceptions of Emotion AI, andits original use of speculative design workshops: While prior work has studied perceptions of a more general public about Emotion AI (Roemmich and Andalibi 2021; Grill and Andalibi 2022), and ethical, historical, and philosophical aspects of Emotion AI, this proposal 'studies up' by focusing on technology (1) workers and (2) students. Emotion AI is rapidly growing and has the potential for transformative impacts, for better or worse; this project will contribute to shaping that potential toward societal benefit.

Broader impacts of this project (Sec. 7), include making AI more inclusive, improved well-being of individuals in society, increased partnerships between academia, industry, and others; and use of science and technology to inform public policy. One of the more unique broader impacts of this proposal comes leveraging feminist technoscience and STS knowledge of observer/observed entanglement. The observation methods of this work, especially interviews and speculative design workshops, will not only observe technology workers and students but also inevitably prompt them to consider the questions posed by interviews and speculative design workshops. Thus, one of the broader impacts of this work is prompting workers and students, who are in positions of relative privilege either working in or studying fields that shape Emotion AI, to consider Emotion AI failures, harms, and mitigation strategies. Through sharing project outputs, the project will invite

[1] These groups are referenced by shorthand later as (1) workers and (2) students.

public informed debate about Emotion AI. Many people are unaware of Emotion AI's existence and growing prevalence; raising awareness is an important and necessary first step toward addressing Browne's call for a *critical biometric consciousness* (Browne 2018). To broaden participation in AI, we must not only recruit more diverse AI practitioners, but also make AI a more inclusive field for diverse people to work. Part of that is making Emotion AI more inclusive in how it models and responds to the diversity, complexity, and cultural nuance of human emotional experiences. This project contributes to that goal with the project output Failure Modes, Harm Vectors, and Strategies for Emotion AI, to help address Emotion AI's cultural bias.

2 Selected Populations and Research Questions

We will compare sociotechnical imaginaries of Emotion AI put forth by (1) workers in technical, design, or management roles who work on Emotion AI systems, features, techniques, or applications in technology companies and (2) undergraduate and graduate students in computer science, AI, design, human-computer interaction, and related fields. These are interesting populations to compare because the AI industry is at a juncture when most current leadership were not subjected to Emotion AI in education or seeking their first job, while in recent years workers and students are increasingly surveilled by Emotion AI in the workplace, education, and job seeking. For example, Humanyze monitors workers for collaborative productivity (*Humanyze* 2021), schools have installed aggression detectors (Gillum and Kao 2019), and an online hiring software withdrew Emotion AI features after an FTC complaint (*EPIC* 2021). These workers and students are impacted by Emotion AI and they are uniquely positioned to change the future of Emotion AI.

2.1 Research Questions

The research questions put corporate sociotechnical imaginaries in comparison with the individual voices of workers and students. Through speculative design, the project not only investigates existing sociotechnical imaginaries but also fosters imagining alternative futures with Emotion AI. By eliciting many different imaginaries and identifying failure modes, the project aims to destabilize Emotion AI's flawed universalizing knowledge claims.

1. (1.1) What sociotechnical imaginaries of Emotion AI are put forth by technology companies? (1.2) How do people who work on Emotion AI in companies imagine futures of Emotion AI? (1.3) How do companies' sociotechnical imaginaries compare to workers' imaginaries? (1.4) How do these imaginaries consider potential failures and harms of Emotion AI? (1.5) How do imaginaries of companies and/or workers compare to those of education programs and students?

2. (2.1) What sociotechnical imaginaries of Emotion AI are put forth by AI education programs? (2.2) How do undergraduate and graduate students in computer science (including AI), design, human-computer interaction, and related fields imagine futures of Emotion AI? (2.3) How do these programs' sociotechnical imaginaries compare to students' imaginaries? (2.4) How do these imaginaries consider failures and harms of Emotion AI? (2.5) How do imaginaries of education programs and students compare to those of companies and workers?

3. As a means of studying imaginaries, how can speculative design workshops advance STS inquiry? What are particular tactics or approaches for better integrating STS concepts and approaches into speculative design research?

As examples, one imaginary from companies could be that Emotion AI will support psychological well-being, to make workers more productive and lower health insurance costs. Individual workers may also value their well-being and productivity while being wary of surveillance limiting their autonomy and influencing insurance rates. Education programs might not specifically address Emotion AI and instead foreground the general societal importance of AI and high-paying job opportunities for future graduates. Individual students might be concerned about Emotion AI's use in virtual testing environments and interview platforms, yet some might be passionate about virtual characters that use Emotion AI to respond to human emotions.

Recent efforts to reform the tech industry have underscored the importance of attending to power dynamics within organizations (Raval et al. 2022). With that in mind, this project studies differences between institutional, company-performed sociotechnical imaginaries of Emotion AI versus imaginaries of individual workers. This project also fosters smaller groups of workers to imagine alternative futures of Emotion AI, and consider ethical risks of Emotion AI, together in a setting where they are not observed by their employers, through the speculative design workshops.

We are also curious about how cultural background might influence imagined futures with Emotion AI. Many Emotion AI companies are based in the US and Europe, yet many of these students are international. For instance, at the PI's institution, undergraduates are 25% international and graduates are 51% international (*GT Diversity* 2022). For instance, immigrant computer science students might be navigating cultural differences in social performance of emotion, while also being implicitly expected by their courses to internalize universalizing notions of AI's ability to symbolically represent emotion—how might such a student imagine futures of Emotion AI and failures and harms of cultural bias in Emotion AI?

3 Related Work

3.1 What Is Emotion AI?

Terminology varies. Roughly, "*Emotion AI* is a subset of artificial intelligence... that measures, understands, simulates, and reacts to human emotions" (Somers 2019). *Affective computing* can be used as a synonym for Emotion AI. This proposal focuses on Emotion Recognition, ways that Emotion AI aims to predict and classify an individual's emotion. Often, these predictions are based on images and videos of faces (e.g., *iMotions Facial Expression Analysis* 2021). Emotion AI can also use other forms of *biodata*: data about people's bodies and behaviors, such as audio, heart rate, skin conductance, or steps (Calvo et al. 2015; Höök 2021). The growing variety of biodata, analysis techniques, knowledge claims, and applications makes Emotion AI a rapidly evolving space.

As Nafus argues in the introduction to *Biosensing Technologies in Everyday Life*, "Focusing on *biosensing* foregrounds the sensors, and therefore the very physical link back to what is being sensed. This connection is important, and can easily be lost in a more data-centric view" (Nafus 2016, p.xiv). Drawing form Nafus, I prefer the less popular term *emotional biosensing* to emphasize emotional phenomena rather than categorical emotion states, and to foreground biodata as coproduced by sensors, humans, and sociocultural context (Howell, Chuang, et al. 2018).

3.2 Ongoing Controversies of Emotion AI

Proponents of Emotion AI claim it offers benefits for societal well-being (R. Picard 1997; El Kaliouby 2020), productivity (Whelan et al. 2018), and security (Jupe and Keatley 2020). Emotion AI applications include wearable devices for psychological well-being (*Ōura Ring* 2021; *Feel: Biomarkers & Digital Therapeutics for Mental Health* 2021; *Bellabeat: Sync your body and mind*

2021; *Limbic.AI Psychological Therapy* 2021; *Pip stress management* 2021; Umematsu et al. 2019), monitoring students (McStay 2020a), preemptive threat detection in schools and airports (Gillum and Kao 2019; *Sound Intelligence Aggression Detection* 2021), evaluating job candidates (Raghavan et al. 2020; Harwell 2019b; *HireVue* 2018; Butcher 2019; *VCV.ai* 2021), assessing workplace stress and productivity (*Humanyze* 2021; Hernandez, Morris, et al. 2011), monitoring driver alertness (Hernandez, McDuff, et al. 2014; *Affectiva Automotive AI* 2021), and criminal investigation tools (Bittle 2020). Many applications psychologically profile people in terms of stigmatizing characteristics related to under-performance, mental illness, or criminal intent (Howell, Chuang, et al. 2018). Emotion AI can augment video surveillance infrastructure and social media photos, amplifying surveillance capabilities.

In one area of debate, while affective computing experts often claim that categories of emotion can universally transcend social and cultural context (e.g., Calvo et al. 2015), HCI design researchers draw from cultural anthropology to critique these claims as culturally reductive, and offer design tactics for affective computing to embrace the nuance, ambiguity, and diversity of emotion as an asset to enrich the design of computational systems (Boehner, DePaula, et al. 2007; Boehner, DePaula, et al. 2005; Boehner 2009; Sengers et al. 2008). I have offered conceptual lenses for design researchers to engage Emotion AI in less reductive ways (Howell, Chuang, et al. 2018).

In another area of debate, critics argue that **Emotion AI poses grave ethical risks.** The AI Now Institute identifies emotion AI as an emerging and urgent concern poised to amplify existing algorithmic racial and gender bias (Crawford et al. 2019). A 2021 report on Emotion AI in China highlights surveillance and civil liberties risks and calls for bans (Marda and Ahmed 2021). Media studies and law scholars have called on policymakers to ban what they call *Physiognomic AI*, attempts to infer human character using AI that includes Emotion AI. They argue these kinds of inferences extend historical practices of physiognomy, racist pseudoscience, and oppression (Stark and Hutson 2021). Psychologists of affect critique shaky scientific foundations (Barrett et al. 2019) and overstated knowledge claims (Chen and Hao 2020) of Emotion AI. Princeton Associate Professor of Computer Science Narayanan, co-author of the book *Fairness and Machine Learning* (Barocas et al. 2019), describes Emotion AI as "fundamentally dubious" "AI snake oil" that does not and fundamentally cannot work (Narayanan 2021).

Emotion AI companies promise improvements in workplace productivity, well-being, healthcare, education, school safety, airport security, and other areas. Companies such as Intel, Google, Amazon, Zoom, Affectiva, iMotions, Feel, Morphcast, Visage, and HireVue offer Emotion AI services (Kaye 2022; *Cloud Vision* 2022; *Amazon Rekognition* 2022; *Zoom IQ* 2022; *Affectiva Media Analytics* 2021; *iMotions* 2022; *Feel: Biomarkers & Digital Therapeutics for Mental Health* 2022; *Morphcast* 2022; *Visage* 2022; *HireVue* 2018). Sometimes, Emotion AI is withdrawn after criticism. For example, HireVue removed psychological predictions of job candidates after an FTC complaint (Harwell 2019b; Harwell 2019a; *HireVue* 2018; *EPIC* 2021). Microsoft significantly curtailed its Emotion AI citing ethical concerns (Vincent 2022). These varied trajectories illustrate how Emotion AI is still in flux. How people imagine futures of Emotion AI—whether desirable, institutional sociotechnical imaginaries that may be used to justify companies working with Emotion AI, or their counterparts of the unethical possibilities for harm put forth by critics of Emotion AI—influences what Emotion AI products or features companies choose to pursue, curtail, or avoid.

Picard, credited with founding the field of affective computing with her 1997 book *Affective Computing* (R. Picard 1997), commented in a 2019 interview that, *"The way that some of this technology is being used... worries me so deeply, that it's causing me to pull back myself on a lot of the things that we could be doing, and try to get the community to think a little bit more about... if we're going to go forward with that, how can we do it in a way that puts forward safeguards that protect people?"* (R. W. Picard 2019).

3.3 Anticipating Emotion AI Failures and Harms

Algorithmic audits of Emotion AI seek to uncover its biases and inaccuracies to anticipate failures. It is helpful to consider how Emotion AI works. Many Emotion AI algorithms rely on facial imagery: First, facial recognition algorithms are used to recognize a face (finding a face in an image, not linking to identity). Then, Emotion AI algorithms look for particular Facial Action Units (FAUs), sub-components of facial expressions such as raised lip corners or a furrowed brow, and associate combinations of FAUs with categorical emotions (*iMotions Facial Expression Analysis* 2021). Of course, even with accurate recognition of particular facial expressions, the link between facial expressions and psychological or emotional meaning is deeply social, cultural, and context-specific—all areas where AI performs poorly.

Evidence of bias with Emotion AI is still emergent, but it seems likely that most Emotion AI algorithms based on facial imagery have racist and ageist bias as well as lower accuracy for people with glasses or face coverings. Racist bias in Emotion AI may stem in part from Emotion AI's reliance on facial recognition, which has well-documented racist bias (Buolamwini and Gebru 2018; Grother et al. 2019). Yet, Emotion AI may not exhibit all the same tendencies as facial recognition; one study found that some FAUs were better recognized for Black faces (Pahl et al. 2022). Disturbingly, another study found that some Emotion AIs categorized Black faces as more angry or contemptuous than white faces (Rhue 2018). Evaluated Emotion AIs performed significantly worse for people over age 34, and for FAUs around the eyes for people wearing glasses (Pahl et al. 2022). Wearing masks or other face coverings would likely prevent recognition of some FAUs and lead to worse Emotion AI performance. So, while Emotion AI is likely racist and ageist, Emotion AI may also present additional axes of marginalization such as whether people wear glasses or cover their face for health, religious, or other reasons.

Systematically documented evidence of harms from Emotion AI are difficult to gather, in part because Emotion AI is usually deployed in tandem with broader systems. Emotion AI debates described above often make arguments in terms of scenarios of what *could* or *might* happen with Emotion AI. These scenarios often draw parallels to other technologies such as facial recognition or recidivism prediction, or trace links to historical practices such as physiognomy. This emphasis on scenarios motivates this project's focus on sociotechnical imaginaries. Because communities most impacted by Emotion AI are not yet known, this project engages workers and students because these groups are increasingly surveilled by Emotion AI in work, school, and hiring, and because these groups are uniquely positioned to improve Emotion AI.

3.4 Studying Preferences and Future Visions of AI

Prior work studied a general public's preferences around Emotion AI. Interviews with US social media users found they had negative perceptions of emotion recognition on social media (Roemmich and Andalibi 2021), and raised concerns around accuracy, privacy, transparency, and contestability (Grill and Andalibi 2022). Survey respondents rated Emotion AI as more acceptable when used for product advertising rather than hiring decisions (Engelmann et al. 2022). A UK survey found people under age 34 rate emotion recognition as more acceptable (McStay 2020b). McStay argues there is an urgent need to define privacy restrictions around Emotion AI because there is currently a weak consensus among industry, legal, policy, and NGO stakeholders regarding privacy for Emotion AI, with different stakeholders motivated by different competing interests (ibid). So, our project's expected outputs (Scenarios, Failure Modes, Harm Vectors, and Strategies) are specifically tailored for policymakers and tech executives to contribute to defining privacy and data protections with Emotion AI.

Figure 1: My design research reimagines dominant techno-narratives. **(left)** I created shirts with embedded biosensors and color-changing biodata displays as an intentionally ambiguous indicator of emotional excitement. The displays consist of subtle e-textile embroidery and can be worn without attracting attention. I invited pairs of friends to wear these garments throughout their daily lives, and interviewed them about their social emotional reflections. This design artifact was not intended to solve a problem or become something people would want to use regularly; rather, the artifact was designed as an elicitation technique, or critical technology probe (Boehner, Vertesi, et al. 2007), to study people's experiences of an alternative emotional biosensing system. Results explored issues of the perceived biopolitical authority of biodata (Howell, Devendorf, Vega Gálvez, et al. 2018; Howell, Devendorf, Tian, et al. 2016). **(right)** This project critiques sociotechnical imaginaries of smart cities (Sadowski and Bendor 2019) that claim data can transparently represent people. I invited pairs of strangers to sit on a bench that amplified their live unfiltered heart sounds, and studied their experiences and reflections on this auditory biodata display (Howell, Niemeyer, et al. 2019). Participants described appreciating the vitality of the stranger sitting next to them while while accepting the opacity, or lack of 'insight' from the biodata. This design explored postcolonial philosopher Glissant's concept of opacity. While the opacity of algorithms has laudably been critiqued elsewhere, Glissantian opacity refers to the opacity of people and the limits of translation and representation, and argues for respecting the inscrutability of human interiority as a prerequisite for upholding the humanity of the Other (Glissant 1997).

Speculative design has been used to investigate future visions of AI. Soden et al. hosted a speculative design workshop to investigate future potential harms of Humanitarian AI (Soden et al. 2022). Framing corporate concept videos as a kind of speculative design, critical discourse analysis on Google Glass and Microsoft Hololens demo videos surfaced sociotechnical imaginaries around these emerging technologies (Wong and Mulligan 2016). Emergent work explores adapting hazards analysis from other engineering fields and bringing it to AI engineering (Martelaro, Smith, et al. 2022). Our project's expected output of Failure Modes, Harm Vectors, and Strategies can be used to guide the kinds of algorthimic audits and system evaluations that should take place with existing Emotion AI deployments to determine whether these harms are occurring, as well as guide the development of future Emotion AI applications in more beneficial directions. This project builds on these speculative design approaches and adapts them for Emotion AI.

4 Relation to PI's Existing Work

My research integrates concepts from STS and cultural anthropology into design research. I use computational, design, and interview research methods to investigate epistemological and biopolitical issues in emotional biosensing. Although much related work uses the terms Emotion AI or affective computing (Sec. 3.1), I frame my work in terms of *emotional biosensing* to foreground ongoing processes of sensing, coproduction, and social performance of emotion and related concepts such as affect and mood (Howell, Chuang, et al. 2018). Drawing from cultural anthropology (Schüll 2016) and feminist new materialism (Barad 2003), I have suggested key conceptual shifts and design tactics for emotional biosensing (Howell, Chuang, et al. 2018) and for diffractive ways of working with biodata more generally (Sanches, Howell, et al. 2022). My work aims to imagine more ethical, inclusive futures with emotional biosensing through speculative design and related approaches in the broader field of design futuring.

These artifacts (Fig. 1) aimed to imagine and experientially explore alternative biodata epistemologies. Yet in studying participants' experiences with these artifacts, I was struck by how steeped participants were in dominant sociotechnical imaginaries; alternatives seemed almost unthinkable. This motivates my current interest in articulating existing imaginaries of Emotion AI as well as hosting speculative design workshops to foster imagining alternatives with Emotion AI.

As I strive to help shift collective imagination toward more ethical, inclusive futures with Emotion AI, my most recent work advances speculative design and design futuring methods. For example, to more deeply explore challenges of using speculative design artifacts to critique sociotechnical imaginaries, I did a retrospective trioethnography around failures in design research (Howell, Desjardins, et al. 2021). My work has contributed reflective lenses for design futuring (Kozubaev et al. 2020), more diverse perspectives for design futuring (Howell, F. Schulte, et al. 2021; Sharma et al. 2022) and fabulations of alternatives ways of living and knowing with biodata (Tsaknaki et al. 2022). Aligning with my commitment to polyvocality in design futuring, and typical of the norms of these HCI venues, these works are frequently coauthored, and in these highly collaborative efforts I am happy to let more junior academics take the lead author spotlight.

I am an assistant professor in the School of Literature, Media, and Communication in the Liberal Arts College at Georgia Tech. In the humanities unit of a technical school, I am uniquely well placed to employ STS and design methods to study sociotechnical imaginaries of Emotion AI, and to hire and train the research team. I also maintain connections with industry, having worked with Google ATAP's Project Jacquard (Devendorf et al. 2016), Google AI's Teachable Machine (Carney, Webster, et al. 2020), Intel, and a startup acquired by Spotify.

5 Project Plan

Research team A PhD student will be recruited to work on this project, with 24 months of graduate research assistant (GRA) funding requested. An undergraduate research assistant (RA) will be recruited to work on this project, with 100 hours requested. The PI is actively recruiting students. The PI is situated within Digital Media, in the School of Literature, Media, and Communication, in the College of Humanities, at Georgia Tech. This highly interdisciplinary program attracts students interested in the social, ethical, and creative implications of AI.

(RQ1.1) Industry, institutional imaginaries: Critical discourse analysis We will use critical discourse analysis to study public materials from Emotion AI companies worldwide in the past 5 years. We will cover companies listed as Emotion AI leaders by market watch websites (e.g., *VentureRadar* 2022), as well as highly ranked companies that offer Emotion AI features (*Forbes* 2022). The selection rationale is that smaller companies focused on Emotion AI may offer a more focused vision of Emotion AI's potential, while highly ranked companies may have more cultural influence on sociotechnical imaginaries even if their Emotion AI offerings are relatively limited. This will likely include Microsoft, Intel, Google, Amazon, Zoom, iMotions, Affectiva/SmartEye, Feel, MorphCast, Visage, etc., because these companies illustrate active, varied trajectories with Emotion AI. **We plan to select about 30 companies worldwide. Data collection:** For these companies, we will gather advertisements, product page text images and videos, industry white papers, and public interviews by company leadership. We will select about 2-3 text and/or visual materials for each company, about 60-90 materials total. We will select on the basis of relevance to Emotion AI, especially applications and ways of living with Emotion AI. For example, a demo video showcasing future ways of living with AI, some of which includes Emotion AI, is more relevant to our analysis than documentation explaining how to use an API for Emotion AI. **Analysis:** To

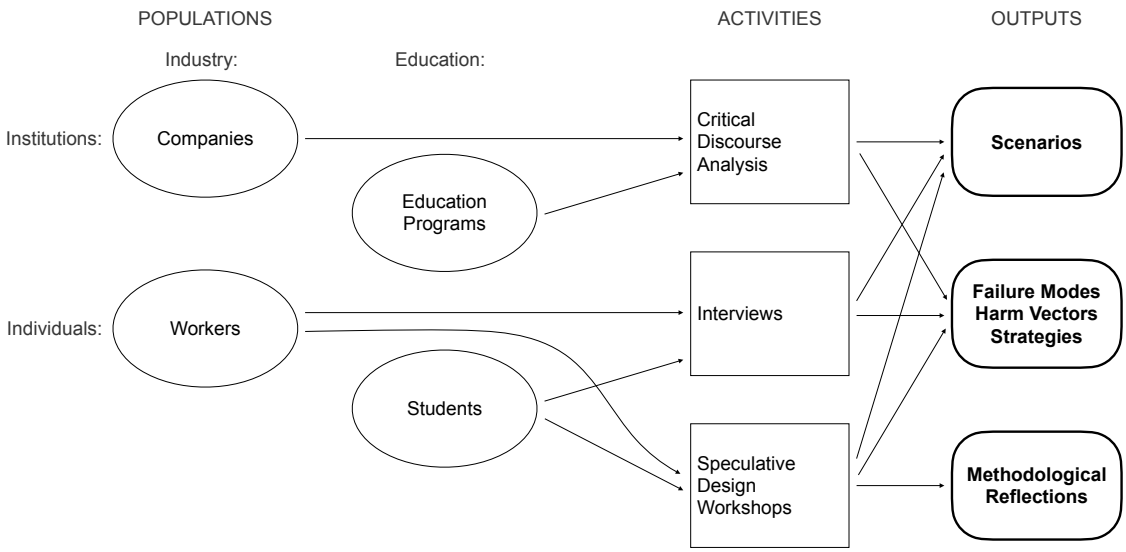

Figure 2: **Project plan:** Critical discourse analysis will study sociotechnical imaginaries of Emotion AI put forth by Emotion AI companies and AI education programs. Interviews and speculative design workshops will study individual and small group imaginaries of Emotion AI from workers and students. These will help develop Scenarios of ways of living with Emotion AI. Throughout interviews and workshops, we will ask workers and students to describe potential failures and harms of Emotion AI, as well as strategies for preventing, mitigating, or addressing these issues, to inform the Failure Modes, Harm Vectors and Strategies. Finally, not shown on the diagram, methodological reflections on hosting the speculative design workshops will be used to identify tactics for using speculative design for STS inquiry.

analyze these materials, for text, we will use critical discourse analysis focusing on how ideas get foregrounded as important, social actors (individual or institutional), social actions (encouraged or discouraged), effects, and settings (Gee 2004). For videos and photos, we will observe people's physical appearance, emotions, relations, actions, effects, and settings (Dyer 2008). We will use qualitative coding (Saldaña 2016) to examine these textual and visual elements, and will also allow unexpected concepts to emerge (Charmaz 2006). The GRA will code all materials; the PI will code an overlapping quarter to reach shared understanding of codes. These methods have been shown to be effective for analyzing sociotechnical visions of emergent technology (Wong and Mulligan 2016).

(RQ2.1) Education, institutional imaginaries: Critical discourse analysis This mirrors the plan for (RQ1.1) except as outlined here. We will use critical discourse analysis to study public materials from AI education programs. We will select from AI education programs listed among the top 100 based on citation rankings (*EduRank* 2021), selecting for programs that mention Emotion AI. If not enough programs mention Emotion AI or synonyms, we will expand this criteria to programs that foreground AI in social contexts (e.g., socially aware robots or chatbots). **We plan to select about 30 education programs worldwide. Data collection:** For these programs, we will examine marketing materials for AI education programs, academic program recruitment materials, university websites, university leadership white papers, websites of hackathons or exhibitions with themes related to AI or Emotion AI, and other public-facing materials.

(RQs 1.2, 1.4) Industry, individual worker imaginaries: Interviews Semi-structured interviews will ask workers involved in Emotion AI products or features in technical, design, user experience, or product management roles to describe their understanding of Emotion AI, what personally motivates them or spurs their interest in Emotion AI, what future applications and ways of living they hope for or fear with Emotion AI (RQ1.2), how they consider potential failures and harms of Emotion AI, and what strategies they can suggest (if any) for mitigating, preventing, or

redressing these issues. This progression of questioning gradually moves from introductory questions to topics that might demand more self-reflection or criticality. To broach potentially sensitive topics such as failures of Emotion AI or potential harms, we will initially pose these questions through the lens of engineering and design practices of identifying issues for iterative improvement (e.g., Sharp and Macklin 2019). We will also draw from emergent work showing the potential of hazards analysis for AI engineering (Martelaro, Smith, et al. 2022). Workers and students in engineering and design roles around Emotion AI are uniquely well positioned to identify not all but many key failure modes of Emotion AI that can have important ramifications for Emotion AI's potential benefits and harms. After broaching the topic of failure with smaller, tractable problems, we will prompt interviewees to expand their consideration to include failures and potential harms that may not have solutions on the horizon. The PI previously worked as a professional software developer, and recently conducted trioethnographic research on failures in design research (Howell, Desjardins, et al. 2021)—good background experience for designing a semi-structured interview guide that asks about failures of Emotion AI.

Each interview will last about one hour, and **the research team will interview about about 30 workers**. For the first few interviews, the GRA and PI will conduct interviews together, to help the GRA refine their interview technique. Then, the GRA will conduct the remaining interviews solo. **Data collection:** Interviews will be conducted by video call or Zoom, audio recorded, and transcribed. **Analysis:** Interview transcripts will be qualitatively coded for emergent themes (Saldaña 2016). The GRA will code all materials; the PI will code an overlapping quarter to reach shared understanding of codes. While allowing unexpected themes to emerge, we will pay particular attention to how participants describe benefits, harms, hopes, fears, and ways of living with Emotion AI, as well as actors, actions, and values. **Recruitment:** Participants will be recruited via the PI's professional network, conferences attended by industry professionals, ML/AI meetup groups, and cold emails to Emotion AI companies targeting those companies selected for critical discourse analysis. The PI is uniquely well positioned for this recruitment, with an undergraduate alumni network from tight-knit, industry-oriented Olin College of Engineering (*Olin* 2022), graduate alumni network that sends many students into Silicon Valley jobs (*UC Berkeley School of Information* 2022), and current position at Georgia Tech.

(RQs 2.2, 2.4) Education, individual student imaginaries: Interviews This mirrors the plan for (RQ1.2) except as outlined here. **We plan to interview about about 30 students**. **Recruitment:** We will forward our recruitment call to contacts at the universities selected for critical discourse analysis, and other universities where the PI has connections. The PI has a letter of collaboration from computing faculty in CMU HCII, computing faculty at National University of Singapore, and from a computing postdoc at University of the Arts London—they have agreed to forward study recruitment calls. The PI was enmeshed in HCI and CS organizations during their PhD at Berkeley. The PI has numerous co-authors worldwide. These collaborators will be asked to forward study recruitment calls.

(RQ1.2) Industry, small group worker imaginaries: Speculative design workshops Speculative design workshops will convene small groups of workers from different companies to imagine many alternative futures with Emotion AI, and potential benefits and harms of those futures. We will first briefly introduce Emotion AI, with beneficial and harmful examples, and ask participants to introduce themselves and how they think about Emotion AI. Then we will use the *Timelines* activity, which asks participants to imagine future news headlines, both positive and negative, about an emergent technology, as well as fictional future social media posts from different stakeholder positions (Wong and Nguyen 2021). We choose this activity because it is designed to encourage wide-ranging ideas about alternative futures with emergent technology, avoid

utopian or dystopian extremes, require participants to consider perspectives other than their own, and support values advocacy for industry workers (ibid); this aligns with our project's interest in imagining future ways of living with Emotion AI. Next, we will play the card game *What Could Possibly Go Wrong?*, as a playful and creative way for participants to generate many possible failure modes and harm vectors. This card game was developed for helping automotive UI professionals imagine harm vectors (Martelaro and Ju 2020); with Martelaro's permission we plan to adapt the deck for Emotion AI. Finally, we will ask each participant to individually free write for 5-10 minutes about an Emotion AI scenario of interest to them. **Each workshop session will have about 15 participants in 3 groups of 5 and last about 2 hours, and we will host 3 workshop sessions, for about 45 participants.** By engaging small groups, we aim to generate many imagined futures of Emotion AI beyond individuals' pre-existing notions. By doing this with workers from different companies away from their employers, we hope to enable them to express ideas different from or even perhaps critical of their employer's mission. By having structured group activities, we aim to inspire participants' thinking and avoid having a few loud voices dominate the conversation. By having both group and individual activities, we aim to have group discussions spur participants' thinking, but still hear from each person's perspective. **Data collection:** We will audio record and photograph the workshops. Audio will be transcribed. We will photograph and transcribe written materials produced by the *Timelines* activity and the individual free-write. We will photograph the arrangements resulting from the What Could Possibly Go Wrong? card game. **Analysis:** Textual data will be qualitatively coded for emergent themes (Saldaña 2016), especially attending to hopes, fears, benefits, harms, actors, actions, and values around Emotion AI. The visual layouts of materials in the group activities carry meaning about how different text phrases relate to one another; we will consider these linkages in analyzing the transcriptions of these materials. **Recruitment:** We will use the same recruitment tactics as for the interviews with workers and invite interviewees back for workshops if they are interested. Also, to facilitate gaining access to relatively large groups of busy industry professionals, we will partner with existing industry public meetup groups to host a special meetup session that includes the workshop activities. For example, we have a letter of collaboration from Carney, who works at Google AI and hosts a monthly Machine Learning User Experience meetup (Carney and H 2022).

(RQ2.2) Education, small group student imaginaries: Speculative design workshops This mirrors the plan for (RQ1.2) except as outlined here. **Each workshop session will have about 15 participants in 3 groups of 5 and last about 2 hours, and we will host 3 workshop sessions, for about 45 participants.** Each workshop will be at a different university, for students within that university. **Recruitment:** The two faculty letters of collaboration have agreed to forward workshop recruitment calls, and are considering partnering to host workshops at their institutions. The PI has strong ties to their alma mater UC Berkeley and can recruit students through various mailing lists. The PI can host one workshop at Georgia Tech, their current institution. The PI plans to proactively network to host workshops at other universities. The workshops are during the second year of the grant, so there is planning time. The workshop can be hosted as a standalone event or as a special session of a course or seminar series.

(RQs 1.3, 1.5, 2.3, 2.5) Comparing imaginaries Critical discourse analysis, interviews, and workshops all study what actors (whether individual or organizational), actions and effects (whether desirable or undesirable), and settings or contexts are imagined for Emotion AI. This describes scenarios of future ways of living with Emotion AI. Comparison is a key mode of analysis for studying sociotechnical imaginaries (Jasanoff and Kim 2015). We understand sociotechnical imaginaries as broader than individual scenarios of specific kinds of actors, actions, effects, and settings. We will use the individual scenarios surfaced by the above activities to articulate sociotechnical imaginaries

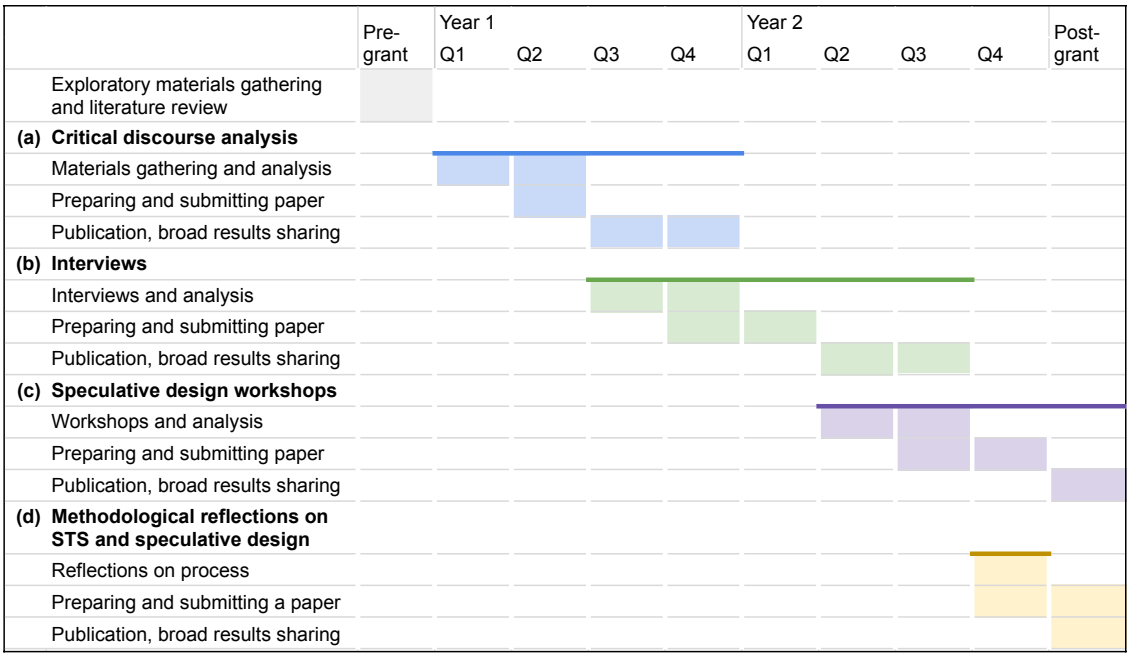

Figure 3: Preliminary materials gathering and literature review has begun. Critical discourse analysis will examine (1) industry marketing and PR materials around Emotion AI and (2) university institutional framings of AI and Emotion AI to address the first parts of RQ1 and RQ2. Interviews will ask about 30 workers and about 30 students how they imagine futures of Emotion AI to address the second parts of RQ1 and RQ2. Speculative design workshops will engage about 45 workers and about 45 students in imagining Emotion AI futures to more expansively engage the second parts of RQ1 and RQ2. Throughout, we will inquire how these imaginaries consider ethical risks of Emotion AI, especially of misclassification and marginalization, to address the third parts of RQ1 and RQ2. Finally, first-person methodological reflections on the process of conducting the speculative design workshops will address RQ3.

as broader descriptions of ways of living with Emotion AI that encompass many potential scenarios. We will aim to articulate sociotechnical imaginaries of industry, sociotechnical imaginaries of education programs, and imaginaries of workers and students, and compare these to one another. We might find that a common sociotechnical imaginary of Emotion AI is shared between industry and education programs, that they have competing sociotechnical imaginaries, that imaginaries are too divergent to comprise a sociotechnical imaginary at all, or something else. Any such finding still contributes to understanding the imagination(s) motivating Emotion AI debates.

(RQ3 & Project Output) Methodological reflections on STS and speculative design

We will reflect on how the workshops combined STS and speculative design methods. The GRA and PI will each write daily field notes during the speculative design workshops phase of the project. Each week, GRA and PI will each write a reflective memo about their experiences and thought process working working to integrate STS concepts and speculative design methods. After conducting the workshops, GRA and PI will interview one another about these field notes and memos, elaborating on points of synergy and friction between STS and speculative design. These interviews will be transcribed and, alongside field notes and reflective memos, qualitatively coded for emergent themes. The PI did a similar duoethnographic process for reflecting on methodological issues in design research (Howell, Desjardins, et al. 2021). Through this, we will reflect on what worked well and what did not work well in terms of the workshop design and theoretical engagements with STS and speculative design. From this, we will provide suggestions for future work engaging sociotechnical imaginaries and speculative design.

IRB Some research activities will require IRB approval. There is sufficient time between the funding decision date and project start date to obtain IRB approval.

(Project Output) Emotion AI Scenarios The research activities above will elicit the actors, actions, effects, settings, and whether these effects are considered desirable or undesirable, or beneficial or harmful, to whom. This comprises a scenario. We will also know the provenance of the scenario (i.e., was it articulated by a company, education program, worker, or student?). The GRA will summarize each scenario in a 1-3 sentences, and create a database associating each scenario with its attributes and provenance. The RA will present this database as an engaging website that allows browsing scenarios as well as searching and filtering by attributes.

(Project Output) Failure Modes, Harm Vectors, Strategies Emotion AI can fail in unique ways and cause unique algorithm harms due to the sociocultural complexity of emotion. For engineers, this project output provide problems to work to prevent or mitigate via technical advances. For designers, it will provide problems to work to prevent or redress through designing new Human-AI collaborative interactions, where in-situ human understanding of emotion can supplant or correct an Emotion AI mistake. For policymakers, it provides areas of focus for enacting protections around data sharing and use. Identified failure modes and harm vectors also support *algorithmic contestability*, the ability to contest algorithmic knowledge claims, which is crucial for people negatively impacted by algorithmic decision making and their advocates (Vaccaro et al. 2019; Hirsch et al. 2017). This project's emphasis on harm vectors also responds to critiques of AI ethical guidelines for lack of attention to societal impacts (Stark and Hoffmann 2019) by highlighting harm vectors as negative social impacts to avoid.

Failure modes and harm vectors will be identified in interviews and workshops by asking participants to describe ways they think Emotion AI can fail and/or potentially cause harm, and strategies for prevention, mitigation, or redress. In analyzing the data collected during interviews and workshops, the GRA will form emergent clusters of commonly articulated failure modes and harm vectors and summarize any strategies for prevention, mitigation, or redress articulated by participants. The GRA will summarize failure modes, harm vectors, and strategies in 1-3 sentences each. Failure modes outline ways that Emotion AI can malfunction; these errors could be innocuous depending on context. As an example of a failure mode, "The Emotion AI system detects an emotion and assumes it is related to the present context, but the emotion is not related to the present context." For instance, classifying smiles as happy could be inaccurate when smiling is meant only as politeness in some contexts. Vectors of harm typically depend on how the Emotion AI system is used. As an example of a vector of harm, "The Emotion AI detected stress in a criminal suspect. The algorithmic classification of stress is used as evidence of guilt. There are other reasons the person felt stressed, such as the interrogation setting or a recent death in the family" (e.g., Bittle 2020). As an example strategy, "Avoid using Emotion AI predictions as evidence in court." We acknowledge that strategies suggested by workers and students may be limited, and we plan to share this output with AI policy and ethics experts to invite their critiques and additions. The RA will present this information as a website modeled after PlatformAbuse.org (Zajac et al. 2021), which outlines known potential for abuse on social platforms alongside mitigation strategies.

Sharing project outputs Conference presentations will be posted freely online. Publications will be open access, and the PI will write summaries explaining the paper for a general audience and post to Medium, their research website, and share via university communications. The project output Scenarios will be shared as a searchable website created by the RA. The GRA and PI will generate lesson plans and discussion prompts for the scenarios to be used as AI literacy education materials. The RA will present the project output Failure Modes, Harm Vectors, and Strategies as a website modeled after PlatformAbuse.org (Zajac et al. 2021) that outlines known harm vectors

and mitigation strategies. The research team will compile each output into an executive summary one-pager targeted at product managers, tech leadership, and policymakers. We will share papers, paper summaries, websites, and one-pagers back with worker and student participants.

6 Intellectual Merit

This proposal's unique focus on sociotechnical imaginaries of Emotion AI with workers and students will offer much-needed empirical insights for decision-making in AI policy and curriculum. Prior work has focused on how laypeople consider Emotion AI (e.g., Grill and Andalibi 2022; Roemmich and Andalibi 2021). While including the voices of these data subjects is important, it is also important to understand how technology workers and students consider Emotion AI. Their perspectives may or may not align with dominant sociotechnical imaginaries of in the technology industry. Workers and students in Emotion AI-related fields are increasingly surveilled in work, school, job seeking, transit, and healthcare, and workers and students are uniquely influential on the development of Emotion AI. Prior work engaging policymakers, industry leaders, and laypeople about Emotion AI suggests a critical, limited timeframe for reaching consensus on appropriate privacy standards and data protections for Emotion AI (McStay 2020b). This shows a need for work of this project contributing to shaping the growth of Emotion AI toward desirable, beneficial futures and away from harms.

7 Broader Impacts

Understanding sociotechnical imaginaries of Emotion AI matters because articulating these imaginaries can inform trajectories of Emotion AI in terms of industry development and policy regulations. The time for this research is now because Emotion AI is rapidly growing and evolving, for better or worse. The project's anticipated outputs—a taxonomy of desirable and undesirable Scenarios with Emotion AI, and Failure Modes, Harm Vectors, and Strategies for Emotion AI—can help shape the development of Emotion AI.

Broadening participation in AI Participant recruitment will target women and other underrepresented minorities for speculative design workshops imagining future ways of living with Emotion AI. These workshops will engage people in critical discussions around AI without technical gatekeeping, a noted strength of combining speculative design and STS (Vertesi et al. 2019). Techniques underlying Emotion AI have documented bias against people with darker skin tones and women (e.g., Buolamwini and Gebru 2018), so including these populations in our recruitment is essential. Also, the burden of improving Emotion AI should not fall to those most impacted by its biases; rather, anticipating and preventing harms of Emotion AI is everyone's moral responsibility, and especially the responsibility of those in positions of privilege who can take more direct action to change Emotion AI—especially workers and students in AI and related fields.

To broaden participation in AI, AI must not only recruit more diverse practitioners, the field of AI must also become more inclusive in how it represents and analyzes diverse populations, especially for socioculturally complex phenomena such as emotion. This project invites AI workers and students to imagine more inclusive, ethical Emotion AI futures that anticipate and mitigate failures and harms of Emotion AI that, at present, are all too common. The project's expected output of Failure Modes, Harm Vectors, and Strategies for Emotion AI can help developers and designers of Emotion AI systems work to address and prevent exclusionary harms of Emotion AI, thus making Emotion AI a more inclusive field in which to work.

Improved STEM education There is increasing interest in AI literacy and AI ethics in education (e.g., Long and Magerko 2020). This project's engagement with students via interviews and workshops offers an informal educational intervention in AI literacy and AI ethics education for Emotion AI. The project output Scenarios can serve as educational material, with lesson plans and discussion prompts we will create.

Increased public scientific literacy and public engagement with science and technology Speculative design can offer a way for broad audiences to engage in discussions about emerging science and technology without requiring technical expertise (Vertesi et al. 2019). The project output Scenarios will be broadly distributed via not only academic papers and presentations but also an engaging website and social media posts targeting a general public.

Improved well-being of individuals in society Emotion AI has shown promise in supporting mental health through emotion regulation (e.g., Slovak et al. 2022). Yet, Emotion AI posts grave ethical risks (Crawford et al. 2019). The project outputs of Scenarios, Failure Modes, Harm, Vectors, and Strategies can be use by policymakers, technologists, and designers, in addressing, preventing, or mitigating Emotion AI harms and advancing beneficial societal impacts.

Increased partnerships between academia, industry, and others The project will engage technology workers and students at different companies to study their perspectives on Emotion AI. Findings and project outputs will be shared back with participants in academia and industry and can lead to ongoing partnerships.

Increased economic competitiveness of the U.S. While the US was initially at the forefront of Emotion AI development, it has since fallen behind. Emotion AI was spearheaded by Picard at the MIT Media Lab in the late 1990s and early 2000s, with startups Affectiva (founded 2009) and Empatica (founded 2013) pushing the bounds of physiological sensing and Emotion AI analyses. Since then, some of Affectiva's core Emotion AI technology went to German-based company iMotions, and now both iMotions and Affectiva are subsidiaries of Sweden-based SmartEye. Empatica no longer focuses on Emotion AI. Today, most major players in Emotion AI are based abroad, with US companies such as Microsoft, Google, and Intel lagging behind as they incrementally roll out Emotion AI capabilities that offer a fraction of the functionality of leading companies. While Emotion AI is projected to nearly double to USD 43B between 2022-2027 (*Markets* 2022), the US is now behind the curve on Emotion AI. This project's outputs of Scenarios of Emotion AI can help strategically position US companies to develop effective Emotion AI applications and help the US regain its competitive economic advantage in Emotion AI. The project output Failure Modes, Harm Vectors, and Strategies will help avoid not only harms but also PR disasters that could curtail Emotion AI's development in the US. This project includes extensive engagement with industry practitioners, and project outputs will be shared with industry practitioners for ongoing dialogue.

Improved national security China is leading the development of expansive Emotion AI surveillance (Marda and Ahmed 2021). The US is is lagging behind on Emotion AI development. The project output Scenarios can help inform the development of future Emotion AI applications for US national security while balancing surveillance with civil liberties.

Use of science and technology to inform public policy The expected outputs of this project can inform public policy on Emotion AI. The websites of Scenarios and Failure Modes, Harm Vectors, and Strategies, as well as one page summaries, will be shared with AI policy experts.

8 Results from Prior NSF Support

The PI has not previously received funding from the NSF.

References Cited

Affectiva Automotive AI (2021). en. URL: http://go.affectiva.com/auto (visited on 07/16/2021).

Affectiva Media Analytics (2021). URL: https://go.affectiva.com/affdex-for-market-research (visited on 07/19/2021).

Amazon Rekognition (2022). *Amazon Rekognition API Documentation: Emotion*. URL: https://docs.aws.amazon.com/rekognition/latest/APIReference/API_Emotion.html (visited on 06/22/2022).

Barad, Karen (2003). "Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter". In: *Signs* 28.3, pp. 801–831. ISSN: 0097-9740. DOI: 10.1086/345321.

Barocas, Solon, Moritz Hardt, and Arvind Narayanan (2019). *Fairness and Machine Learning*. fairmlbook.org.

Barrett, Lisa Feldman, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak (July 2019). "Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements". en. In: *Psychological Science in the Public Interest* 20.1, pp. 1–68. ISSN: 1529-1006, 1539-6053. DOI: 10.1177/1529100619832930. URL: http://journals.sagepub.com/doi/10.1177/1529100619832930 (visited on 12/13/2019).

Bellabeat: Sync your body and mind (Jan. 2021). en-US. URL: https://bellabeat.com/ (visited on 02/02/2021).

Bittle, Jake (Mar. 2020). "Lie detectors have always been suspect. AI has made the problem worse." en. In: *MIT Technology Review*. URL: https://www.technologyreview.com/2020/03/13/905323/ai-lie-detectors-polygraph-silent-talker-iborderctrl-converus-neuroid/ (visited on 06/21/2021).

Boehner, Kirsten (Nov. 2009). "Reflections on representation as response". en. In: *interactions* 16.6, p. 28. ISSN: 10725520. DOI: 10.1145/1620693.1620700. URL: http://portal.acm.org/citation.cfm?doid=1620693.1620700 (visited on 05/02/2017).

Boehner, Kirsten, Rogério DePaula, Paul Dourish, and Phoebe Sengers (2005). "Affect: from information to interaction". en. In: *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility (CC'05)*. ACM Press, pp. 59–68. ISBN: 978-1-59593-203-7. DOI: 10.1145/1094562.1094570. URL: http://portal.acm.org/citation.cfm?doid=1094562.1094570 (visited on 01/12/2016).

— (Apr. 2007). "How emotion is made and measured". en. In: *International Journal of Human-Computer Studies* 65.4, pp. 275–291. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2006.11.016. URL: http://linkinghub.elsevier.com/retrieve/pii/S1071581906001844 (visited on 01/12/2016).

Boehner, Kirsten, Janet Vertesi, Phoebe Sengers, and Paul Dourish (2007). "How HCI Interprets the Probes". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ISBN: 978-1-59593-593-9. DOI: 10.1145/1240624.1240789.

Bowker, Geoffrey C. and Susan Leigh Star (1999). *Sorting Things out: Classification and Its Consequences*. Inside Technology. Cambridge, Mass: MIT Press. ISBN: 978-0-262-02461-7.

Browne, Simone (2018). "B®anding Blackness - Biometric Technology and the Surveillance of Blackness". In: *Sondra Perry: Typhoon Coming On*. Ed. by Sondra Perry and Amira Gad. Köln: Walther König.

Buolamwini, J. and T. Gebru (2018). "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification". In: *Proceedings of Machine Learning Research*. Vol. 81, pp. 77–91.

Butcher, Mike (2019). *The robot-recruiter is coming — VCV's AI will read your face in a job interview*. en-US. URL: http://social.techcrunch.com/2019/04/23/the-robot-recruiter-is-coming-vcvs-ai-will-read-your-face-in-a-job-interview/ (visited on 12/13/2019).

Calvo, Rafael, Sidney D'Mello, Jonathan Gratch, and Arvid Kappas, eds. (Jan. 2015). *The Oxford Handbook of Affective Computing*. Oxford University Press. ISBN: 978-0-19-994223-7. URL: http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199942237.001.0001/oxfordhb-9780199942237 (visited on 05/18/2016).

Carney, Michelle and Nikki H (2022). *Machine Learning & User Experience (MLUX) (San Francisco, CA) Meetup*. URL: https://www.meetup.com/mluxmeetup/ (visited on 07/20/2022).

Carney, Michelle, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen (Apr. 2020). "Teachable Machine: Approachable Web-Based Tool for Exploring Machine Learning Classification". en. In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, pp. 1–8. ISBN: 978-1-4503-6819-3. DOI: 10.1145/3334480.3382839. URL: https://dl.acm.org/doi/10.1145/3334480.3382839 (visited on 12/21/2020).

Charmaz, Kathy (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Pine Forge Press.

Chen, Angela and Karen Hao (Feb. 2020). "Emotion AI researchers say overblown claims give their work a bad name. A lack of government regulation isn't just bad for consumers. It's bad for the field, too." en. In: *MIT Technology Review*. URL: https://www.technologyreview.com/2020/02/14/844765/ai-emotion-recognition-affective-computing-hirevue-regulation-ethics/ (visited on 07/19/2021).

Crawford, Kate, Roel Dobbe, Theodora Dryer, Genevieve Fried, Ben Green, Elizabeth Kaziunas, Amba Kak, Varoon Mathur, Erin McElroy, Andrea Nill Sánchez, Deborah Raji, Joy Lisi Rankin, Rashida Richardson, Jason Schultz, Sarah Myers West, and Meredith Whittaker (2019). *AI Now 2019 Report*. Tech. rep. New York: AI Now Institute, p. 100. URL: https://ainowinstitute.org/AI_Now_2019_Report.pdf.

Devendorf, Laura, Joanne Lo, Noura Howell, Jung Lin Lee, Nan-Wei Gong, M. Emre Karagozler, Shiho Fukuhara, Ivan Poupyrev, Eric Paulos, and Kimiko Ryokai (May 2016). ""I don't Want to Wear a Screen": Probing Perceptions of and Possibilities for Dynamic Displays on Clothing". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY, USA: Association for Computing Machinery, pp. 6028–6039. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858192. URL: https://doi.org/10.1145/2858036.2858192 (visited on 12/22/2020).

Dyer, Gillian (Mar. 2008). *Advertising as Communication*. Zeroth. Routledge. ISBN: 978-1-134-97234-0. DOI: 10.4324/9780203158340. URL: https://www.taylorfrancis.com/books/9781134972340 (visited on 07/18/2022).

EduRank (Aug. 2021). URL: https://edurank.org/cs/ai/ (visited on 07/20/2022).

El Kaliouby, Rana (2020). *Girl Decoded: A Scientist's Quest to Reclaim Our Humanity by Bringing Emotional Intelligence to Technology*. Random House.

Morphcast (2022). *Emotion AI Interactive Video Platform*. URL: https://www.morphcast.com/ (visited on 06/22/2022).

Engelmann, Severin, Chiara Ullstein, Orestis Papakyriakopoulos, and Jens Grossklags (June 2022). "What People Think AI Should Infer From Faces". In: *2022 ACM Conference on Fairness, Accountability, and Transparency*. Seoul Republic of Korea: ACM, pp. 128–141. ISBN: 978-1-4503-9352-2. DOI: 10.1145/3531146.3533080. URL: https://dl.acm.org/doi/10.1145/3531146.3533080 (visited on 06/27/2022).

EPIC (2021). *EPIC - HireVue, Facing FTC Complaint From EPIC, Halts Use of Facial Recognition*. https://epic.org/2021/01/hirevue-facing-ftc-complaint-f.html.

Visage (2022). *Face Analysis: Detect detailed data for people's gender, age and emotions and build engaging experiences*. URL: https://visagetechnologies.com/face-analysis/ (visited on 06/22/2022).

Feel: Biomarkers & Digital Therapeutics for Mental Health (2021). en. URL: https://www.myfeel.co/ (visited on 02/02/2021).

Feel: Biomarkers & Digital Therapeutics for Mental Health (2022). URL: https://www.myfeel.co/ (visited on 03/21/2022).

Gee, James Paul (2004). "Discourse Analysis: What Makes It Critical?" In: *An Introduction to Critical Discourse Analysis in Education*. Ed. by Rebecca Rogers. Hoboken: Lawrence Erlbaum Associates. ISBN: 978-1-4106-0978-6.

GT Diversity (2022). *Georgia Institute of Technology - Main Campus Diversity: Racial Demographics & Other Stats*. URL: https://www.collegefactual.com/colleges/georgia-institute-of-technology-main-campus/student-life/diversity/ (visited on 06/22/2022).

Gillum, Jack and Jeff Kao (June 2019). "Aggression Detectors: The Unproven, Invasive Surveillance Technology Schools Are Using to Monitor Students". en. In: *ProPublica*. URL: https://features.propublica.org/aggression-detector/the-unproven-invasive-surveillance-technology-schools-are-using-to-monitor-students/ (visited on 07/19/2021).

Glissant, Édouard (1997). "For Opacity". In: *Poetics of Relation*. Trans. by Betsy Wing. The University of Michigan Press, pp. 189–194.

Grill, Gabriel and Nazanin Andalibi (Apr. 2022). "Attitudes and Folk Theories of Data Subjects on Transparency and Accuracy in Emotion Recognition". In: *Proceedings of the ACM on Human-Computer Interaction*. DOI: 10.1145/3512925. URL: https://dl.acm.org/doi/abs/10.1145/3512925 (visited on 06/27/2022).

Grother, Patrick, Mei Ngan, and Kayee Hanaoka (2019). *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*. Tech. rep. NISTIR 8280. National Institute of Standards and Technology, U.S. Department of Commerce. URL: https://doi.org/10.6028/NIST.IR.8280.

Harwell, Drew (Nov. 2019a). "A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job". In: *Washington Post*. ISSN: 0190-8286.

— (Nov. 2019b). "Rights group files federal complaint against AI-hiring firm HireVue, citing 'unfair and deceptive' practices". en-US. In: *Washington Post*. ISSN: 0190-8286. URL: https://www.washingtonpost.com/technology/2019/11/06/prominent-rights-group-files-federal-complaint-against-ai-hiring-firm-hirevue-citing-unfair-deceptive-practices/ (visited on 07/19/2021).

Hernandez, Javier, Daniel McDuff, Xavier Benavides, Judith Amores, Pattie Maes, and Rosalind Picard (June 2014). "AutoEmotive: bringing empathy to the driving experience to manage stress". In: *Proceedings of the 2014 companion publication on Designing interactive systems*. DIS Companion '14. New York, NY, USA: Association for Computing Machinery, pp. 53–56. ISBN: 978-1-4503-2903-3. DOI: 10.1145/2598784.2602780. URL: https://doi.org/10.1145/2598784.2602780 (visited on 07/16/2021).

Hernandez, Javier, Rob R. Morris, and Rosalind Picard (2011). "Call Center Stress Recognition with Person-specific Models". In: *Proceedings of the 4th International Conference on Affective Computing and Intelligent Interaction*. Vol. 1. ACII'11. Berlin, Heidelberg: Springer-Verlag, pp. 125–134. ISBN: 978-3-642-24599-2. URL: http://dl.acm.org/citation.cfm?id=2062780.2062798 (visited on 01/12/2016).

HireVue (2018). en. URL: https://www.hirevue.com/products/video-interviewing (visited on 12/25/2018).

Hirsch, Tad, Kritzia Merced, Shrikanth Narayanan, Zac E. Imel, and David C. Atkins (June 2017). "Designing Contestability: Interaction Design, Machine Learning, and Mental Health". In: *Proceedings of the 2017 Conference on Designing Interactive Systems*. DIS '17. New York, NY, USA: Association for Computing Machinery, pp. 95–99. ISBN: 978-1-4503-4922-2. DOI: 10.1145/3064663.3064703. URL: https://doi.org/10.1145/3064663.3064703 (visited on 07/08/2021).

Höök, Kristina (2021). *Affective Computing*. en. URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/affective-computing (visited on 07/16/2021).

Howell, Noura, John Chuang, Abigail De Kosnik, Greg Niemeyer, and Kimiko Ryokai (Nov. 2018). "Emotional Biosensing: Exploring Critical Alternatives". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW, 69:1–69:25. DOI: 10.1145/3274338. URL: https://doi.org/10.1145/3274338 (visited on 12/22/2020).

Howell, Noura, Audrey Desjardins, and Sarah Fox (Nov. 2021). "Cracks in the Success Narrative: Rethinking Failure in Design Research through a Retrospective Trioethnography". In: *ACM Transactions on Computer-Human Interaction* 28.6, 42:1–42:31. ISSN: 1073-0516. DOI: 10.1145/3462447.

Howell, Noura, Laura Devendorf, Rundong (Kevin) Tian, Tomás Vega Gálvez, Nan-Wei Gong, Ivan Poupyrev, Eric Paulos, and Kimiko Ryokai (June 2016). "Biosignals as Social Cues: Ambiguity and Emotional Interpretation in Social Displays of Skin Conductance". In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. DIS '16. New York, NY, USA: Association for Computing Machinery, pp. 865–870. ISBN: 978-1-4503-4031-1. DOI: 10.1145/2901790.2901850. URL: https://doi.org/10.1145/2901790.2901850 (visited on 12/22/2020).

Howell, Noura, Laura Devendorf, Tomás Alfonso Vega Gálvez, Rundong Tian, and Kimiko Ryokai (Apr. 2018). "Tensions of Data-Driven Reflection: A Case Study of Real-Time Emotional Biosensing". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: Association for Computing Machinery, pp. 1–13. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3174005. URL: https://doi.org/10.1145/3173574.3174005 (visited on 12/22/2020).

Howell, Noura, Britta F. Schulte, Amy Twigger Holroyd, Rocío Fatás Arana, Sumita Sharma, and Grace Eden (May 2021). "Calling for a Plurality of Perspectives on Design Futuring: An Un-Manifesto". In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. New York, NY, USA: Association for Computing Machinery, pp. 1–10. ISBN: 978-1-4503-8095-9. DOI: 10.1145/3411763.3450364. URL: https://doi.org/10.1145/3411763.3450364 (visited on 05/09/2021).

Howell, Noura, Greg Niemeyer, and Kimiko Ryokai (May 2019). "Life-Affirming Biosensing in Public: Sounding Heartbeats on a Red Bench". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: Association for Computing Machinery, pp. 1–16. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300910. URL: https://doi.org/10.1145/3290605.3300910 (visited on 12/22/2020).

Humanyze (2021). *Humanyze: Workplace Analytics Software: Unlock Your Business Potential*. en-US. URL: https://humanyze.com/ (visited on 07/16/2021).

iMotions Facial Expression Analysis (2021). URL: https://imotions.com/biosensor/fea-facial-expression-analysis/ (visited on 08/10/2021).

iMotions (2022). *iMotions: Unpack Human Behavior*. URL: https://imotions.com/ (visited on 06/22/2022).

Jasanoff, Sheila and Sang-Hyun Kim (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. en. University of Chicago Press. DOI: 10.7208/chicago/9780226276663.001.0001.

Jupe, Louise Marie and David Adam Keatley (Dec. 2020). "Airport Artificial Intelligence Can Detect Deception: Or Am i Lying?" In: *Security Journal* 33.4, pp. 622–635. ISSN: 0955-1662, 1743-4645. DOI: 10.1057/s41284-019-00204-7.

Kaye, Kate (Apr. 2022). *Intel Thinks Its AI Knows What Students Think and Feel in Class*. URL: https://www.protocol.com/enterprise/emotion-ai-school-intel-edutech (visited on 06/22/2022).

Kozubaev, Sandjar, Chris Elsden, Noura Howell, Marie Louise Juul Søndergaard, Nick Merrill, Britta Schulte, and Richmond Y. Wong (Apr. 2020). "Expanding Modes of Reflection in Design Futuring". en. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, pp. 1–15. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376526.

Limbic.AI Psychological Therapy (2021). *Limbic.AI is enabling the best psychological therapy. Our AI-powered patient reporting tool makes it easier than ever for patients to provide their clinician with the data they need to treat optimally.* URL: https://limbic-website.netlify.app// (visited on 07/16/2021).

Long, Duri and Brian Magerko (Apr. 2020). "What is AI Literacy? Competencies and Design Considerations". en. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, pp. 1–16. ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376727. URL: https://dl.acm.org/doi/10.1145/3313831.3376727 (visited on 08/28/2020).

Marda, Vidushi and Shazeda Ahmed (2021). *Emotional Entanglement: China's emotion recognition market and its implications for human rights*. Tech. rep. Article 19. URL: https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf.

Markets (2022). URL: https://www.marketsandmarkets.com/Market-Reports/emotion-detection-recognition-market-23376176.html (visited on 07/04/2022).

Martelaro, Nikolas and Wendy Ju (Sept. 2020). "What Could Go Wrong? Exploring the Downsides of Autonomous Vehicles". In: *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '20. New York, NY, USA: Association for Computing Machinery, pp. 99–101. ISBN: 978-1-4503-8066-9. DOI: 10.1145/3409251.3411734. URL: https://doi.org/10.1145/3409251.3411734 (visited on 07/19/2022).

Martelaro, Nikolas, Carol J. Smith, and Tamara Zilovic (Mar. 2022). *Exploring Opportunities in Usable Hazard Analysis Processes for AI Engineering*. DOI: 10.48550/arXiv.2203.15628. arXiv: 2203.15628 [cs]. URL: http://arxiv.org/abs/2203.15628 (visited on 07/19/2022).

McStay, Andrew (July 2020a). "Emotional AI and EdTech: serving the public good?" In: *Learning, Media and Technology* 45.3. Publisher: Routledge _eprint: https://doi.org/10.1080/17439884.2020.1686016, pp. 270–283. ISSN: 1743-9884. DOI: 10.1080/17439884.2020.1686016. URL: https://doi.org/10.1080/17439884.2020.1686016 (visited on 07/19/2021).

— (Jan. 2020b). "Emotional AI, Soft Biometrics and the Surveillance of Emotional Life: An Unusual Consensus on Privacy". In: *Big Data & Society* 7.1, p. 2053951720904386. ISSN: 2053-9517. DOI: 10.1177/2053951720904386.

Merrill, Nick, Richmond Wong, Noura Howell, Luke Stark, Lucian Leahu, and Dawn Nafus (2017). "Interrogating Biosensing in Everyday Life". In: *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*. DIS '17 Companion. New York, NY, USA: ACM, pp. 364–367. ISBN: 978-1-4503-4991-8. DOI: 10.1145/3064857.3064865. URL: http://doi.acm.org/10.1145/3064857.3064865.

Nafus, Dawn (2016). "Introduction". In: *Quantified: Biosensing Technologies in Everyday Life*. Ed. by Dawn Nafus. MIT Press.

Narayanan, Arvind (Jan. 2021). *How to recognize AI snake oil*. Princeton, New Jersey. URL: https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf (visited on 09/05/2021).

Noble, Safiya Umoja (2018). *Algorithms of oppression: how search engines reinforce racism*. New York: New York University Press.

Olin (2022). *Olin College of Engineering*. URL: https://www.olin.edu/ (visited on 07/20/2022).

Ōura Ring (2021). *Ōura Ring: Accurate Health Information Accessible to Everyone*. en. URL: https://ouraring.com (visited on 07/07/2021).

Pahl, Jaspar, Ines Rieger, Anna Möller, Thomas Wittenberg, and Ute Schmid (June 2022). "Female, White, 27? Bias Evaluation on Data and Algorithms for Affect Recognition in Faces". In: *2022 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '22. New York, NY, USA: Association for Computing Machinery, pp. 973–987. ISBN: 978-1-4503-9352-2. DOI: 10.1145/3531146.3533159. URL: https://doi.org/10.1145/3531146.3533159 (visited on 06/27/2022).

Picard, Rosalind (1997). *Affective Computing*. The MIT Press.

— (2022). *Projects Page*. URL: https://www.media.mit.edu/people/picard/projects/ (visited on 07/04/2022).

Picard, Rosalind W. (June 2019). *Interview with Rosalind Picard on June 17, 2019: Affective Computing, Emotion, Privacy, and Health*. Date: June 17, 2019. URL: https://www.media.mit.edu/articles/rosalind-picard-affective-computing-emotion-privacy-and-health-artificial-intelligence-podcast/ (visited on 06/27/2021).

Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy (Jan. 2020). "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices". In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. arXiv: 1906.09208, pp. 469–481. DOI: 10.1145/3351095.3372828. URL: http://arxiv.org/abs/1906.09208 (visited on 09/05/2021).

Raval, Noopur, Rida Qadri, Richmond Y. Wong, Tamara Kneese, and Alex Hanna (Apr. 2022). "Considerations for Building Solidarity among Academic and Tech Workers: Thinking through Access, Positionality and Limits to Collective Action". In: *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. New Orleans LA USA: ACM, pp. 1–3. ISBN: 978-1-4503-9156-6. DOI: 10.1145/3491101.3516511. URL: https://dl.acm.org/doi/10.1145/3491101.3516511 (visited on 06/27/2022).

Rhue, Lauren (2018). "Racial Influence on Automated Perceptions of Emotions". en. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.3281765. URL: https://www.ssrn.com/abstract=3281765 (visited on 12/13/2019).

Roemmich, Kat and Nazanin Andalibi (Oct. 2021). "Data Subjects' Conceptualizations of and Attitudes Toward Automatic Emotion Recognition-Enabled Wellbeing Interventions on Social Media". In: *Proceedings of the ACM on Human-Computer Interaction*. DOI: 10.1145/3476049. URL: https://dl.acm.org/doi/abs/10.1145/3476049 (visited on 06/27/2022).

Sadowski, Jathan and Roy Bendor (May 2019). "Selling Smartness: Corporate Narratives and the Smart City as a Sociotechnical Imaginary". In: *Science, Technology, & Human Values* 44.3, pp. 540–563. ISSN: 0162-2439. DOI: 10.1177/0162243918806061.

Saldaña, Johnny (2016). *The Coding Manual for Qualitative Researchers*. 3E [Third edition]. Los Angeles ; London: SAGE. ISBN: 978-1-4739-0248-0 978-1-4739-0249-7.

Sanches, Pedro, Noura Howell, Vasiliki Tsaknaki, Tom Jenkins, and Karey Helms (Apr. 2022). "Diffraction-in-Action: Designerly Explorations of Agential Realism Through Lived Data". In: *CHI Conference on Human Factors in Computing Systems*. CHI '22. New York, NY, USA: As-

sociation for Computing Machinery, pp. 1–18. ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3502029.

Sanches, Pedro, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas (May 2019). "HCI and Affective Health: Taking Stock of a Decade of Studies and Charting Future Research Directions". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: Association for Computing Machinery, pp. 1–17. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300475. URL: https://doi.org/10.1145/3290605.3300475 (visited on 01/22/2021).

Sas, Corina, Kristina Höök, Gavin Doherty, Pedro Sanches, Tim Leufkens, and Joyce Westerink (July 2020). "Mental Wellbeing: Future Agenda Drawing from Design, HCI and Big Data". In: *Companion Publication of the 2020 ACM Designing Interactive Systems Conference*. DIS' 20 Companion. New York, NY, USA: Association for Computing Machinery, pp. 425–428. ISBN: 978-1-4503-7987-8. DOI: 10.1145/3393914.3395920. URL: https://doi.org/10.1145/3393914.3395920 (visited on 02/02/2021).

Schüll, Natasha Dow (Sept. 2016). "Data for Life: Wearable Technology and the Design of Self-Care". In: *BioSocieties* 11.3, pp. 317–333. ISSN: 1745-8552, 1745-8560. DOI: 10.1057/biosoc.2015.47.

Sengers, Phoebe, Kirsten Boehner, Michael Mateas, and Geri Gay (June 2008). "The disenchantment of affect". en. In: *Personal and Ubiquitous Computing* 12.5, pp. 347–358. ISSN: 1617-4909, 1617-4917. DOI: 10.1007/s00779-007-0161-4. URL: http://link.springer.com/10.1007/s00779-007-0161-4 (visited on 01/06/2019).

Sharma, Sumita, Britta Schulte, Rocío Fatás, Noura Howell, Amy Twigger Holroyd, and Grace Eden (Apr. 2022). "Design Futuring for Love, Friendship, and Kinships: Five Perspectives on Intimacy". In: *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. CHI EA '22. New York, NY, USA: Association for Computing Machinery, pp. 1–14. ISBN: 978-1-4503-9156-6. DOI: 10.1145/3491101.3516388.

Sharp, John and Colleen Macklin (2019). "Failure". In: *Iterate: Ten Lessons in Design and Failure*. Cambridge, MA: The MIT Press, pp. 23–42. ISBN: 978-0-262-03963-5.

Slovak, Petr, Alissa N. Antle, Nikki Theofanopoulou, Claudia Daudén Roquet, James J Gross, and Katherine Isbister (2022). "Designing for Emotion Regulation Interventions: An Agenda for HCI Theory and Research". In: DOI: 10.48550/ARXIV.2204.00118. URL: https://arxiv.org/abs/2204.00118 (visited on 07/04/2022).

Soden, Robert, Aleks Berditchevskaia, Erin Coughlan de Perez, Manveer Kalirai, Shreyasha Paudel, Isabel Stewart, Saurav Poudel, and Sakun Joshi (2022). "What Could Possibly Go Wrong? Speculative Practice Towards Anticipating the Negative Consequences of Humanitarian AI". In: *FAccT CRAFT Sessions: ACM Conference on Fairness, Accountability, and Transparency - Workshops and Panels*. URL: https://facctconference.org/2022/acceptedcraft.html (visited on 06/27/2022).

Somers, Meredith (Mar. 2019). *Emotion AI, Explained*.

Sound Intelligence Aggression Detection (2021). URL: https://www.soundintel.com/products/overview/aggression/ (visited on 09/05/2021).

Stark, Luke (2019). "Affect and Emotion in digitalSTS". In: *DigitalSTS: A Field Guide for Science & Technology Studies*. Ed. by Janet Vertesi and David Ribes. Princeton, New Jersey: Princeton University Press.

Stark, Luke and Anna Lauren Hoffmann (2019). "Data Is the New What? Popular Metaphors & Professional Ethics in Emerging Data Culture". In: *Journal of Cultural Analytics*. DOI: 10.22148/16.036. URL: https://culturalanalytics.org/2019/05/data-is-the-new-what-

popular-metaphors-professional-ethics-in-emerging-data-culture-2/ (visited on 11/01/2019).

Stark, Luke and Jevan Hutson (Sept. 2021). *Physiognomic Artificial Intelligence*. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.3927300. URL: https://papers.ssrn.com/abstract=3927300 (visited on 06/20/2022).

Pip stress management (2021). *The Pip teaches you how to manage your stress better. It allows you to see your stress levels, connecting your emotions with engaging apps, teaching you not only how to recognize stress, but to know a life without it.* en-gb. URL: https://thepip.com/en-gb/how-it-works/ (visited on 07/16/2021).

Forbes (2022). *Top 100 Digital Companies*. URL: https://www.forbes.com/top-digital-companies/list/ (visited on 07/18/2022).

VentureRadar (2022). *Top Emotion AI Companies*. URL: https://www.ventureradar.com/keyword/Emotion%20AI (visited on 07/18/2022).

Cloud Vision (2022). *Try It! Cloud Vision API Demo*. URL: https://cloud.google.com/vision/docs/drag-and-drop (visited on 06/22/2022).

Tsaknaki, Vasiliki, Pedro Sanches, Tom Jenkins, Noura Howell, Boer Laurens, and Afroditi Bitzouni (2022). "Fabulating Biodata Futures for Living and Knowing Together". In: *Designing Interactive Systems*.

UC Berkeley School of Information (2022). URL: https://www.ischool.berkeley.edu/ (visited on 07/20/2022).

Umematsu, Terumi, Akane Sano, and Rosalind Picard (July 2019). "Daytime Data and LSTM can Forecast Tomorrow's Stress, Health, and Happiness". In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. ISSN: 1558-4615, pp. 2186–2190. DOI: 10.1109/EMBC.2019.8856862.

Vaccaro, Kristen, Karrie Karahalios, Deirdre K. Mulligan, Daniel Kluttz, and Tad Hirsch (Nov. 2019). "Contestability in Algorithmic Systems". In: *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. CSCW '19. New York, NY, USA: Association for Computing Machinery, pp. 523–527. ISBN: 978-1-4503-6692-2. DOI: 10.1145/3311957.3359435. URL: https://doi.org/10.1145/3311957.3359435 (visited on 07/08/2021).

VCV.ai (2021). *VCV Online Recruitment Automation Software: Virtual Staffing Solutions*. URL: https://vcv.ai (visited on 07/19/2021).

Vertesi, Janet, David Ribes, Laura Forlano, Yanni Loukissas, and Marisa Leavitt Cohn (2019). "Engaging, Designing, and Making Digital Systems". In: *DigitalSTS: A Field Guide for Science & Technology Studies*. Ed. by Janet Vertesi and David Ribes. Princeton, New Jersey: Princeton University Press.

Vincent, James (June 2022). *Microsoft to Retire Controversial Facial Recognition Tool That Claims to Identify Emotion*. URL: https://www.theverge.com/2022/6/21/23177016/microsoft-retires-emotion-recognition-azure-ai-tool-api (visited on 06/21/2022).

Whelan, Eoin, Daniel McDuff, Rob Gleasure, and Jan vom Brocke (Feb. 2018). "How Emotion-Sensing Technology Can Reshape the Workplace". In: *MIT Sloan Management Review*.

Wong, Richmond Y. and Deirdre K. Mulligan (2016). "When a Product Is Still Fictional: Anticipating and Speculating Futures Through Concept Videos". In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. DIS '16. New York, NY, USA: ACM, pp. 121–133. ISBN: 978-1-4503-4031-1. DOI: 10.1145/2901790.2901801. URL: http://doi.acm.org/10.1145/2901790.2901801 (visited on 08/02/2019).

Wong, Richmond Y. and Tonya Nguyen (May 2021). "Timelines: A World-Building Activity for Values Advocacy". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing*

Systems. CHI '21. New York, NY, USA: Association for Computing Machinery, pp. 1–15. ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445447. URL: https://doi.org/10.1145/3411764.3445447 (visited on 05/11/2021).

Zajac, Avi, Ji Su Yoo, Nicole Chi, and Franchesca Spektor (2021). *PlatformAbuse.Org*. URL: https://platformabuse.org (visited on 06/25/2021).

Zoom IQ (2022). https://explore.zoom.us/en/conversational-intelligence/.