The Human in Emotion Recognition on Social Media: Attitudes, Outcomes, Risks

Nazanin Andalibi, Justin Buss
School of Information, University of Michigan
Ann Arbor, Michigan, USA
{andalibi, buss}@umich.edu

ABSTRACT

Emotion recognition algorithms recognize, infer, and harvest emotions using data sources such as social media behavior, streaming service use, voice, facial expressions, and biometrics in ways often opaque to the people providing these data. People's attitudes towards emotion recognition and the harms and outcomes they associate with it are important yet unknown. Focusing on social media, we interviewed 13 adult U.S. social media users to fill this gap. We find that people view emotions as insights to behavior, prone to manipulation, intimate, vulnerable, and complex. Many find emotion recognition invasive and scary, associating it with autonomy and control loss. We identify two categories of emotion recognition's risks: *individual* and *societal*. We discuss findings' implications for algorithmic accountability and argue for considering emotion data as sensitive. Using a Science and Technology Studies lens, we advocate that technology users should be considered as a *relevant social group* in emotion recognition advancements.

Author Keywords
Emotion Recognition; Emotion AI; Social Media; Ethics; Privacy; Fairness; Algorithmic Accountability; AI Ethics

CSS Concepts
Human-centered computing → Empirical studies in HCI

INTRODUCTION

"*Because it's emotional data. It's like therapist notes*," said a participant to us in discussing emotions. Another said, "*Your emotions are so personal…so human.*" Emotions are powerful, mediate human experiences with their surroundings, and impact decision-making and attention [30,40,76,94] online and off. Privacy and emotion are related in many ways; emotions are crucial in users' sense of privacy [112]. Online and off, emotions are often deemed private; Sharing and signaling them to others can be beneficial (e.g., finding support and community, improved wellbeing [10,57,93]), but involve privacy calculations and complex decision-making processes [7,9,16,100].

Emotion Recognition and Emotion Artificial Intelligence (AI) detect and infer emotional states [84]. Despite the deeply personal nature of emotions, AI algorithms are built to recognize, infer, and harvest emotions using data sources such as social media behavior, streaming service use, voice, facial expressions, biometrics, and body language in ways often unknown to users [29,69,84,92]. Such inferences can be used for curating social media news feeds, advertising, and other algorithmic decision-making and manipulation of media environments [54,84,88]. These include applications in many domains such as market research, customer service, and advertising; healthcare and wellbeing; employment; the workplace; entertainment; the automotive industry; education; politics; interactive systems; law enforcement; and surveillance [36,84,118]. Interest in emotion recognition spans industry, academia, and government.

Companies like Google, Facebook, Amazon, Snapchat, Spotify, and IBM either already use emotion recognition or have filed patents (e.g., [1,19,26,56,68,81,84,90]). Many technology start-ups focus on emotion recognition (e.g., [65,67,73,84,102]). Additionally, a growing body of research in fields such as computing, economics, medical informatics, public health, and psychology not only seeks to detect and predict people's emotional and mental states (e.g., depression) from direct expressions (e.g., a social media post or voice command saying "I'm depressed") [29,37] but also from more indirect and obscure expressions (e.g., an Instagram image with no explicit depression-related expression or one's voice features) [79,98]. Lastly, governments can both regulate and use emotion recognition technologies. For example, during the Sochi Olympics, Russian officials used video and emotion analytics to identify agitated attendees by measuring facial muscle vibrations [61]. The emerging emotion recognition market is expected to grow from $123 million in 2017 to $3.8 billion by 2025 [60,118]. It is estimated to reach its plateau in the next 2 to 5 years, and is in the early mainstream phase of market penetration (5-20% of target audience) [36]. The increasing availability of and access to large amounts of data, cheaper computational power, and improved deep learning, natural language processing (NLP), and computer vision techniques have facilitated these shifts [84].

CHI 2020, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04...$15.00.
DOI: https://doi.org/10.1145/3313831.3376680

Prior work provides valuable insights on attitudes towards online privacy (e.g., [46,96]) and algorithms (e.g., [35,59]), and ethics and values in AI and research (e.g., [15,39,74,108,125–127,129,131]). It also provides preliminary insights into how technologists think about emotion recognition [84]. The key perspective missing from these debates is that of the *humans* who produce the data that make emotion recognition possible, and whose experiences are shaped by these technologies – this paper's focus.

Beliefs about emotions and their use matter as they shape our interactions with the world [49], online and off. Due to emotions' roles and sensitivity, this paper focuses on emotions and users' attitudes towards emotion recognition technologies to contribute to our knowledge about socially and ethically responsible use and treatment of data in algorithmic decision-making that impacts humans' personal lives. A first exploratory step is to investigate people's attitudes, values, and reactions in relation to emotion recognition as an emerging technology, as well as the risks and harms *they* perceive and anticipate, short and long term. Focusing on social media and through interviews (N=13) with social media users, we reveal the concerns and attitudes of people whose data make emotion recognition technologies possible and who are influenced by emotion-related algorithmic decision-making in relation to this technology.

Contributions. First, we contribute an account of how people's values and views towards emotions inform their attitudes towards emotion recognition. Participants viewed emotions as a unique data type different from other personal data. They remarked that emotions provide unique insights to behavior and are prone to manipulation; and are intimate, personal, vulnerable, complex, and hard to define. Second, we contribute an understanding of social media users' attitudes towards emotion recognition on these platforms. Participants had varied, but often negative reactions to emotion recognition using social media data. While for some *what* the recognition was used for informed their discomfort, for others it did not. Third, we highlight outcomes and risks related to emotion recognition's use on social media as perceived by participants. The main anticipated outcome was less use, more vague posts, and stopping use. While people may not do what they say they will do, these insights highlights their *values* towards emotion recognition (this study's focus). We also identify perceived risks associated with emotion recognition: 1) *individual* risks related to control, manipulation, exploitation; unfair harm distribution; negative mental health impacts; identity misrepresentation including beyond one's lifetime; and challenges with holding algorithms responsible; and 2) *societal* risks related to social and political control and manipulation. Assessing these findings' generalizability is an area for future work.

At this paper's core, by taking emotion recognition and social media as a context, we foreground the *humans* affected by emerging technologies whose data make these technologies possible. Drawing from Social Constructivism [95] we argue that technology users should be considered as a *relevant social group* [105] in emotion recognition advancements. In situating our findings within broader scholarships, the risks and impacts we uncovered are dimensions of algorithmic accountability (e.g., societal impacts and potential harms) [3,23,52,110]. We argue that technologies that feel into people's emotions should acknowledge people's complexities. Acknowledging this complexity does not necessarily mean building more accurate technologies to infer those complexities; rather, we argue that these innovations must prioritize the preferences and values of the *humans* they impact.

PRIOR WORK

Emotion Recognition. "Emotion recognition" (i.e., interpreting data to decipher one's emotional state) is the first step in making emerging "Emotion AI" technologies possible [36]. Emotion AI involves algorithms that recognize and classify emotions to respond in a "personalized" way [84]. We define "emotional data" as data from which emotional states can be inferred; such data can include emotional content directly (e.g., someone writing they are sad) [7], or indirectly (e.g., a black and white picture one took while feeling sad) [9]. The roots of technologies that gauge emotions go back to the mid to late 1800s as reviewed in [84], but perhaps the most influential in computing is Affective Computing in the 1990s [21]. Emotion researchers across disciplines view emotions from several theoretical perspectives [22]. Much of the emotion recognition research in Computing draws on Ekman [41,42], who identified six "basic" and "universal" emotions: anger, disgust, fear, joy, sadness, and surprise. Some have critiqued the notion of only six emotions [101] and their universality [13,53], and argued that there is no scientific evidence "that a person's emotional state can be readily inferred from his or her facial movements," which others disagree with [14,111].

Companies, People, and Emotion Recognition. U.K.-focused research by McStay [84] is most closely related to this paper. He conducted 100 interviews with stakeholders (e.g., technologists, regulators) and held a workshop with members of these groups. This work identified these stakeholders' views, future directions, and challenges of emotion recognition, and developed ethical guidelines for emotion recognition around autonomy, consent, control, empowerment, freedom, transparency, and trust. To develop these guidelines, the workshops and interviews did *not* address the perspectives of the *people* whose data make emotion recognition possible and who are influenced by it.

McStay [84] further conducted a brief U.K. national survey of attitudes towards Emotion AI and emotion recognition. The survey included five questions about how comfortable U.K. citizens were with emotion recognition in six domains including social media. Results showed that 50.6% were uncomfortable, 30.6% were comfortable if anonymized and not personally linked to them, 8.2% were comfortable even if inferences are linked to them, and 10.4% were not sure.

However, these may be substantially different in the U.S. (and outside the U.K. more broadly) due to different regulations and privacy norms, as well as political ideologies that inform privacy attitudes [120]. The work reviewed here provides key initial insights into people's comfort with emotion recognition, and calls for an *in-depth* understanding of people's attitudes towards emotion recognition, which we address in this study. We extend this prior work by broadening existing research about attitudes towards emotion recognition to a context outside the U.K., deepening past research in this space by identifying the *reasons* behind attitudes and perceived associated outcomes and harms, and including *technology users* as an important stakeholder not included in developing existing guidelines [84] to lay the groundward for future guidelines and work that include users' values and concerns. Our goal is not to address a binary of "should emotion recognition exist or not," but to investigate users' attitudes, values, and concerns around it.

Attitudes Towards Algorithmic Decision-Making. Broadly, support for AI development is mixed among Americans and greater among those who are wealthy, educated, male, or technology experts [35]. Specific to social media, Pew [109] found that users' comfort level with social media companies using their data depends on *how* it is used, covering four contexts in decreasing order of comfort: recommend events, recommend people, show ads, and show messages from political campaigns. Pew's questions and contexts did not cover emotions, or wellbeing (closely linked to emotions [93]) in any way. Pew's respondents likely did not think of emotions when assessing their comfort with algorithms, as many do not know emotional inference is possible (as also reflected in participants' accounts in this paper). We ask, how might these attitudes change if *emotional* data are harvested, or if emotion recognition is used for algorithmic decision-making? Is "how" emotional data is used one or the only concern, or is there more to it?

Emotion Recognition on Social Media. People use social media and express their emotions on them for reasons including identity expression, support exchange, and finding community [6,7,43]. Increasingly, research computationally detects, predicts, and recognizes direct disclosures of emotions and emotional states in social media (e.g., [2,28,29,71,103,117,124]). A recent review of research on mental health state prediction from social media data suggests that much of this work simplifies humans to impersonal "users" or "subjects" [25]. Social media platforms have explored ways to support those in emotional distress using Emotion AI [88] – critiqued for ethical and transparency reasons in opinion (non-empirical) pieces [11] and similarly applauded for possibly being helpful [107]. More recently, academics and technologists have used machine learning and facial recognition to detect emotions and mental health status based on more obscure signs (e.g. visual markers of depression on Instagram [79,98]. New AI systems or methods can cause new privacy harms as inferences can be made even about things people do *not* disclose [34,119]. Yet, potentially helpful applications based on emotion recognition include building agents who provide emotional support to users [70] or helping individuals on the autism spectrum with communication [66,75].

Scholars have examined social media users' attitudes toward using Twitter data to monitor depression for *research*, finding that people were more comfortable with aggregate level monitoring than individual assessments, and concerned about consent, permanence, and privacy [32,33,87]. Researchers have also examined social media users' understanding of and attitudes towards *research* use of *publicly* shared data, finding that contextual factors (e.g., study topic, aggregate vs. individual analysis) matter, and users do not always know what happens to their data [47]. Studies and social media corporate experiments interacting with people's private information (e.g. the "emotion contagion" study [72]) have received public and scholarly attention, and sometimes backlash (e.g., [12,38,55,64,85,86,97,113,130,131]). These studies highlight that people do have concerns around the treatment of their social media data and social media experiences.

Emotion recognition can be used to deliver personalized content (e.g., ads, newsfeed) [84]. For example, it can be used on social media to curate feeds [54,82,84] with negative mental health impacts [54]. Additionally, harvesting emotional data online is part of a "behavioral turn" in digital commerce [89]. "Surveillance capitalism" [132] relies on behavioral manipulation [89] to "provide the right message at the right time to the right person." [84]. Overall, some data uses are perceived as "creepy" [106,121], and data related to health, location, web browsing, age, finances, and private communication are identified as highly sensitive [17,77,78,123]. Research has also used design fiction to examine ethical and privacy implications of emerging technologies [15,39,74,108,125–127] with one focusing on emotions and ads [108] presenting a design for an AI that detects a user's emotional state to promote ads, identifying a "grey area" in technology ethics and data use [108].

In summary, some data are more sensitive than others, and some data uses can be less accepted than others, including on social media. Emotion recognition provides an excellent example of an emerging technology applied to social media that can impact humans in deep, personal ways; People's attitudes towards emotion recognition on social media and the risks and outcomes that they associate with it remains unknown – which we address in this study.

METHODS, DATA, AND ANALYSIS

We conducted semi-structured interviews (N=13) with adult social media users in the U.S. We detail our process next.

Recruitment. We recruited interview participants via a screening survey. We shared the survey on our personal social media accounts which was widely shared by our networks. We also shared it on Craigslist in Detroit, MI and Houston, TX in order to reach a larger and diverse audience.

Research suggests Craigslist as a platform to reach diverse research participants [128]. We chose these cities as two of the most diverse cities in the U.S. [83]. The screening survey was re-shared at least 18 times on Facebook and retweeted 45 times publicly. Impression count for tweets (not *re*tweets) was 11276, more than 6x our follower count. Three participants were acquaintances of one author, in which case the other author conducted interviews to add social distance. The screening survey received 100 responses. Of those responses, we contacted 20 respondents and conducted interviews with those who followed up (*N*=13). Participants received $30, and the study was approved by our IRB.

Screening survey. The screening survey asked respondents if they used social media, if they were located in the U.S., and their age. If one of these criteria was not met (if they responded no to either question or were younger than 18), the survey ended. We asked respondents what social media they used and which ones they posted to regularly. The survey also asked about positive and negative personal experiences from the past year. If respondents had these experiences, they were asked if and where they posted about them on social media. Additionally, the survey included questions about demographics such as race, gender, and education level.

Interview participants. We invited interview participants purposefully based on responses to the survey and collected data – recruitment, data collection, and primary analysis were iterative. Specifically, we contacted respondents who reported experiencing both positive and negative emotional experiences in the past year and *posted* about those on social media. This was because we wanted the participants to have *real* experiences we could ground the interviews in. Positive experiences included getting a new job, an educational accomplishment, or buying a house. Negative experiences included political events, losing a job, the end of a relationship, and (physical and mental) health complications. The one-year time limit ensured reasonable recall about the experiences and social media landscape. All reported posting on at least one platform at least once per week. We also considered age, gender, education, and race to cover a diverse range of experiences as much as possible. Table 1 includes participant information.

We conducted interviews via video or phone call based on the participant's preference, recorded and transcribed the audio, and took notes. Interviews lasted from 77 to 120 minutes (average=106 min).

Limitations and reflections. First, we asked participants about experiences from the past year for higher recall, but there may have been limitations in their recall. Yet, our goal was to examine how people reconstruct meaning and associate values with emotion recognition. Therefore, possible recall issues did not interfere with our goals. Second, several participants shared pre-existing privacy concerns (not on emotions), manifested in adjusting settings or tailoring feeds; however, they had all *still* chosen to post about personal experiences. That said, this may have led to self-selection bias. Third, in line with phenomenological research, our goal was not generalizability and our sample was not representative [104]; For instance, while our sample included typically underrepresented genders, it included fewer men, who may be less willing to discuss emotions [18]. Our sample included five people of color. People of color may be impacted by emerging technology in more harmful ways [99] and their voices are less represented in technology discourse and research. Most participants had some college education or may have been more familiar with technology than an average person. This is common in studies of emerging technology [4,59]. It is crucial to uncover the attitudes of less educated individuals, diverse genders, children, older adults, and people in diverse parts of the world -- important areas for future work. Future work may also evaluate our findings with representative samples and at a large scale.

	Age	**Gender**	**Race**	**Education**	**Social Media**
P1	24	Agender	White	College	FB, TW, RD, TB
P2	58	Woman	White	Graduate	FB, TW, LI
P3	20	Genderfluid	Indian	College	FB, IG, TW, TB, AO3
P4	23	Woman	Asian	Graduate	FB, IG, TW, RD
P5	25	Woman	White	College	TW, SC, TB, DC
P6	43	Woman	Black	College	FB, FBG, IG
P7	28	Woman	White	Graduate	FB, FBG, IG, TW, SC, RD, LI
P8	36	Woman	White	Graduate	FB, FBG
P9	24	Woman	Asian	Graduate	IG, TW
P10	27	Genderqueer	Black	Graduate	FB, FBG, IG, TW, SC, RD, TCH, YT
P11	22	Man	White	High School	FB, FBG, TW, SC, RD, TB
P12	52	Woman	White	College	FB, FBG, IG
P13	39	Woman	White	Some College	FB, FBG, IG, TW, SC

Table 1. Participant demographics. Abbreviations for social media sites: Archive of Our Own: AO3, Discord: DC, Facebook: FB, Facebook Groups: FBG, Instagram: IG, LinkedIn: LI, Reddit: RD, Snapchat: SC, Tumblr: TB, Twitch: TCH, Twitter: TW, YouTube: YT

Interview phases and analysis. The first interview phase investigated people's current social media use, how and why they have or have not used social media to share about meaningful emotional experiences, mental models of what happens to their shared emotional and other data, whether they have noticed any changes online after posting emotional content, what emotions mean to them, and their expectation for privacy with any entity they believe may access their information. The second phase focused on gauging people's attitudes, expectations, and values about emotion recognition for which we used scenarios: We asked participants to imagine positive and negative personal experiences (as discussed earlier in the interview) and social media most relevant to that experience. We then asked them to *imagine* a scenario where the social media site they posted on had used computational methods and their data to infer their emotional states, for example at the time of or after posting. Using prompts and follow-up questions as is common in semi-structured interviews, we then explored participants' attitudes and values towards emotion recognition on social media. Follow-up prompt topics that the analysis presented in this paper draws on included: feelings about and reactions to the scenario and reasons for those feelings; how personal awareness of emotion recognition on social media would affect participants; how the scenario matched with their expectations of what already occurs, and what they desired to occur; and what harms or benefits they anticipate such technology would have for them. By allowing flexibility in how participants interpreted scenarios, we uncovered values towards current systems and imagined futures. This choice is informed by work examining privacy values [126] and understandings of news feed algorithms [44]. What people think algorithms (can) do and their related attitudes (our focus) is as key as what algorithms actually do [44,122].

Scenario-like methods are common in HCI to gather reactions to imagined designs and to explore values and attitudes towards technology [5,20,24,62] or to develop theory [8]. Scenarios are useful when participants may not have direct experience with the phenomena being explored [48]. While potential differences in what people say they will do, and what they will do in practice is a relevant critique, research suggests that in emotional settings people behave similarly in "real life" as they respond to scenarios [63]. This study's goal was to examine people's values surrounding an emerging technology that is hard to access and interpret by non-experts, and not what participants will or will not do in practice in reaction to this technology's deployment, making scenarios an excellent tool to utilize.

We analyzed the data using the constant comparative approach [116]. We met frequently during data collection to discuss primary emerging themes and to refine the interview protocol. One author first open coded five interviews. We then discussed each code in detail, refined codes, and grouped them into larger themes. Another five interviews were coded and grouped into the previous themes or emerging ones via a similar process. The remaining interviews were then coded and codes were organized into the existing themes. No new themes emerged in this phase.

RESULTS

We first describe how participants *conceived of* their emotions and emotional data when considering possibilities of emotion recognition on social media. We then discuss *reactions* to emotion recognition and end with *perceived risks* and *outcomes* of emotion recognition on social media.

Perceptions of Emotions and Emotion Recognition

How do social media users conceptualize emotions when considering how they might be algorithmically analyzed? We provide these insights to set the ground for the rest of our findings, illustrating the unique characteristics of emotions and data about them to social media users.

Emotions provide insights into behavior and can be manipulated to impact behavior. Participants viewed emotions as insights into a person's behavior. As P9 put it: "*I do think that in the society we underestimate how our emotions are connected to our actions and behaviors*," and as P6 elaborated: "*I guess because I think that emotions are part of your body's driving force…dictate your behavior…dictate your health*." Participants noted that understanding emotions could lead to controlling individuals. As P5 said: "*Really genuinely knowing how somebody else is feeling is a key insight into their behavior and their thought processes, and again you can control people based on how they're feeling*." Participants were concerned about emotional data *specifically*, because they felt that emotions could be easily manipulated to impact behavior. These conceptions of emotions provided the foundation for participants' beliefs that emotion recognition can exploit and manipulate human emotions and behavior.

Emotions are intimate, personal, and vulnerable. Participants conceived of emotions as intimate and integral to understanding an individual in depth. As P5 said: "*I guess I would say that to know how someone is feeling is the most intimate understanding of a person*," and P6 elaborated: "*I think your emotions tell a lot about who you are*." Participants viewed emotions as very personal, compared to other kinds of data. For instance, P4 said: "*I guess I find it to be more personal, so I guess that's the reason I do not prefer it crossing into professional contexts.*" Because of this intimacy, participants often wanted to keep emotions private or separate from some parts of their lives. Emotions carried with them some vulnerability like that of a journal or therapy session for participants. For example, P7 said: "*I think because emotions are real. I mean, they're vulnerable parts...*" Along the same lines, P3 noted: "*Again, therapist's notes, right? So, there are things you tell your therapist only she can understand. Right? Between you and her.*" These conceptions of emotions provided the grounds for participants' reactions to emotion recognition in terms of its privacy invasiveness and the extent to which it can cause harm because it engages with such vulnerable data about people's lives as we discuss later.

Emotions are complex and hard to define, even for humans. Participants felt that emotions were complex and not always easily understood even by other humans, let alone algorithms. For instance, P1 reflected on emotions' complexity and said: "*It's just such a hard to define experience even for the person feeling it. It just seems weird to me to quantify that in a way that a computer can understand. Because not a lot of people are, they understand it all that well*." P1 further elaborated that: "*It's not just like people are happy or sad or angry. There's a million things in between...*" Finally, participants noted that not only are emotions complex, but they are also individualized and not universal. For instance, P10 said: "*I feel like the experience of having an emotion, then sort of the lifelong experience of understanding it, learning how to deal with it, learning what triggers it, that introspection, are all sort of uniquely human*." P3 echoed this sentiment: "*Everyone feels happy or sad or whatever, but everyone feels it differently. That's a great part of being alive*." Emotions' complexity perceived by participants contributed to them questioning whether emotions can be truly recognized by algorithms and non-human agents.

Reactions to Emotion Recognition Based on Social Media Data

We identified reactions to using emotion recognition techniques to detecting emotional states in certain times, and predicting emotional states in the future. We present findings about detection (in a moment) and prediction (in the future) separately as participants conceived of them differently. We see how people's perceptions and attitudes towards emotions (as described earlier) shape reactions to emotion recognition technologies' use on social media. Most participants had negative reactions, but a minority were not so uncomfortable.

Reactions to Emotion Detection on Social Media

Having one's emotions detected based on social media data evoked intense feelings for participants. Overall, participants felt that emotion recognition is invasive, scary, and sometimes unnecessary. Some noted that the perceived lack of meaningful consent led to feeling intrusion. For example, P11 said: "*It's not okay… It's intrusive. It's unwarranted. No one gave the permission. I certainly didn't.*" For others, these feelings were about losing the type of control they felt they had in communicating with other humans, when considering an algorithm wanting to understand their emotions like a human would; P3 said: "*For me that's still disturbing, still makes me uneasy. Because I feel it's like your social media trying to understand you like a person. It's trying to be another person to understand you. I don't like that because I don't have any control over that*." People have some degree of control and opportunity to correct misunderstandings in human to human communication in-person. This becomes challenging when they may not know how their human audiences perceive them or their emotions on social media, and even more challenging when algorithms are thought of as the audience. When algorithms read people's emotions, participants felt that they had even less control over how they are read and understood. This reaction was exacerbated by participants' conceptions of emotions as complex and hard to understand, compared to other data types.

Others compared having one's emotions detected to having cameras in one's house; P12 noted: "*If I found out later on that they did that [emotion detection] ... I would be a little upset…because again it's intrusive to me that somebody's doing something and I don't know anything about it…That's why people don't like cameras inside their house. It's like… spies or something are watching you*." The camera in the house metaphor that P12 used is an example of how participants felt about emotion recognition on social media and its implications for their privacy loss – exacerbated by their conceptions of emotions as personal and intimate.

Not only did emotion detection feel like an invasion of privacy, it was also outright scary to some participants. Sometimes this fear was due to algorithms 'seeing into' who one 'really' is; As P2 said: "*It's a little bit scary. It scares me that we're so easily read. It scares me that algorithms can so easily see into who we are*." Other times because participants worried it would enable controlling populations; P5 elaborated on this point: "*Freaked out. I'd wonder what kind of* 1984 *society they're trying to create so they can control the population…That's my paranoia and my English teacher, dystopia brain freaking out but I don't know, lately these days, things are seeming a lot closer to that than anybody would like them to be*." This reaction was exacerbated by thinking of emotions as being prone to control and manipulation to impact emotions and behavior.

Participants also reflected on what this detection may be used for. Some remarked that regardless of the end result of the detection, emotion detection is invasive. For instance, P3 said: "*Good or bad it's still an invasion of privacy*." While other participants felt like it just was not necessary for platforms to be doing emotion detection. For example, P5 said: "*I suppose that would depend on what the purpose was…I guess I'm more inclined, regardless of what the purpose is, even if it's innocent to feel negatively about it. Simply because I don't feel like it's necessary. Why do you need to do that? Don't do that*." Others reflected that not knowing what such a detection will be used for causes concerns, as put by P8: "*If I don't know what they were going to do with that information, I'd be worried*." While for some the personal nature of emotions meant that computationally recognizing them to whatever end is invasive, for others, knowing *what* it was going to be used for mattered.

Reactions to Emotion Prediction on Social Media

Similar to detections, participants were largely uncomfortable with predicting emotional states based on social media data. For some, predictions (about future emotional states) felt like a step further than detections (of current emotions). As P13 put it: "*One thing is to see how I'm feeling at that moment, but to predict how I'm feeling in the future, that's kind of weird*." P3 echoed this sentiment: "*It's one thing for Instagram to have my data or Twitter to have my data and keep it in some server. But for them to be actively reading it in a way that's trying to understand [my future]…I don't like*." Specifically, for P4, predictions were where they drew the line: "*I would be fine with that unless it's like how I would feel in the future..*."

Discomfort was heightened for some because predictions could be made in the first place; some participants believed that emotions were not predictable because even *they* had a hard time predicting *their* own emotions. As P7 said: "*Well, that would be a little weird because I can't even predict how I'm going to feel in the future half the time.*" In this sense, how people related to their own emotions informed how they felt about algorithms relating to their emotions.

For some, similar to emotion detection, what the prediction would be used for did not matter – they still were not comfortable with it for reasons such as lack of control and agency. For example, P7 said: *"I think it's the idea that I'm in control of my emotions and my decisions, and who are you to tell me what I will or won't feel tomorrow, or two days from now, or a year from now*." For others, how this prediction was used did matter. As P2 said: "*How I feel about it depends on how the information is used more than just the* *fact that it's happening. The fact that it's happening doesn't particularly bother me in general…But if it gets into hands where somebody has direct power over me and begins to treat me in a certain way because they believe these predictions, it bothers me a lot*." In thinking about how information gained from emotion recognition will be used, participants reflected on risks associated with emotion recognition – which we discuss later.

Not Concerned About Emotion Recognition on Social Media

Emotion detection and prediction on social media did not always bother participants. A minority (N = 4) were not entirely uncomfortable with it, sometimes because they had accepted it as something that already happens frequently. As P9 described: "*I think we make assumptions as a society all the time, so I don't find this to be disturbing*." Yet, even P9 wanted to know about the process's details: "*I feel like the prediction is not, yeah, it's not bad. Like, I think it's just that we just need to know what they are, like, how that is being constructed and like how that is being coded..*." Other times, participants were not averse to emotion detection and prediction when they perceived their posting behavior on social media to not be controversial. For instance, P12 said: "*I mean, it doesn't really bother me because I don't really get too crazy and radical with anything that I post*." In these cases, the participants had either accepted how decisions are made in the world in general, or used social media to post "non-controversial" content. That said, other participants also employed various privacy protection strategies such as posting non-controversial or vague content, but were *still* uncomfortable with emotion recognition on social media.

Perceived Outcomes and Risks Associated with Emotion Recognition based on Social Media Data

We discuss participants' anticipated *outcomes* of emotion recognition on social media assuming awareness (as opposed to no awareness). These outcomes were about changes in using social media. We then outline the *risks* participants associated with emotion recognition on social media: *individual* and *societal*-level risks.

Outcome: changing social media use. Participants often expressed a sort of "*give and take*" (P5) with general social media use, when not thinking about emotion recognition or knowing that it is a possibility. They often recognized an overall lack of privacy, but at the end of the day it was important for them to use social media because of the benefits they received (e.g., support and community). For instance, P12 said: "*I like Facebook. I don't want to get rid of it. Too many friends on there, and contacts.*" P12 continued: "*I like to keep in touch with my friends…so I'm just careful of what I do, for the most part.*" These insights are echoed by prior work [27].

However, many participants felt that if they were aware that emotion recognition was happening on social media, they would change the way they posted on social media, potentially going as far as ending use altogether. For example, P11 said: "*I would probably just stop using the app and delete my account, because that would creep me out definitely if I knew that was happening. It would make me feel uncomfortable and violated,*" and "*I probably just wouldn't post near as much, if at all. Maybe stop using them because I wouldn't trust them as much.*" Similarly, P5 said: "*I'd get rid of my social media. I'd be like no, I'm done. I don't need this. I can survive without it. Everybody I love, give me your phone numbers and I'll just text you…There's an extent where it's like no way, that's too far*." While some may be able to comfortably leave social media, others rely on it, as we will elaborate in the Discussion.

While some participants already tailored their posts to protect privacy (e.g., posting vague and indirect content), they stated that this tailoring would be more extreme if they were aware that emotion recognition occurred. As P10 put it: "*I'd probably post more vague things, unless it's expressly important. Because once you kind of know a thing exists or an algorithm exists, it's hard not to think about it when posting stuff or, like, try to game it in some way.*" On a similar note, P8 said: "*Then I would be very careful what I post to Facebook… Even though I think I already regulate…*" Examining whether participants *actually* would quit social media or regulate their use in certain ways is not our goal here, neither is possible at this point. Rather, these insights highlight participants' values and concerns, which in practice may or may not lead to changing social media use.

Related to changing social media use and what people would share and not share about themselves, participants also highlighted a tension between what emotion recognition algorithms learn about us and what that means for social media users' identity presentation. P10 discussed how it can be problematic for algorithms to know so much about us because they do not leave any sense of privacy: "*then they [algorithms] pretty much know your whole life rather than sort of the persona that exists online. They have an idea of who you actually are, and that could be a problem*." P10 continued that in response to such a world, people can turn to masking parts of themselves to protect some aspects of their privacy and identity: "*Or, everyone online becomes super fake, so that person is just their persona and no one knows who you actually are, which is just as problematic…*" Using emotion recognition on social media (assuming users' awareness) that can really 'see into' people's emotions, people may share less and less of what is truly meaningful to them and their identities. We return to this in the Discussion.

Individual risk: control, manipulation, and exploitation. Participants were concerned about emotion recognition-enabled controlling of emotions. For instance, P5 said: "*people… don't want to be controlled… If you know how a general population is feeling, you can control the information that's coming out better in a more tactically intelligent way."* Participants, such as P2, also noted that emotion recognition has the potential to manipulate people's views: "*I don't like the idea of being swayed… manipulated… But a very realistic part of me believes that,*

that happens every day in every aspect of our lives. It could be marketers and companies are doing that all the time. Whether it's on social media or not social media. I think the impact is stronger on social media. It bothers me."

Participants considered emotion recognition's use for control and manipulation particularly harmful in bad actors' hands. As P3 said:"*[Emotion recognition] could be used as a way to exploit me, if it gets in the wrong hands. If someone knows I feel really happy or really positive about this kind of content, they can send me an email or something about that kind of content. And then hack into my life...*" P5 echoed similar concerns around using emotion recognition for malicious intent: "*I suppose at first glance, it seems innocent. What are you going to do with the knowledge of a person's emotions, but in the same respect, because they do have the ability to recommend content to you, if they wanted to use theirs for their own nefarious purposes, they could gear certain content towards you based upon the emotions you're having*." This, to participants meant that their emotions would be manipulated and their behaviors impacted.

Participants mentioned particular domains for emotion recognition applications on social media, highlighting the importance of *what* emotion recognition is used for. As P8 said: "*It depends on how they used it…I don't think it would harm me just sitting here knowing that's happening, but I think depending on how they use it, it could harm me*." For example, the potential for emotions being controlled or manipulated enabled through emotion recognition can have remarkable impacts in the marketing domain. As P9 said: "*I think it's risky that companies capitalize on our emotions to sell us products... Product advertisements that are based on our emotions are harmful because most of the [consumers] already don't understand what they're doing with their emotions... and then buying into the advertisements…If those two go together, it's like living without thinking.*" The potential for emotion recognition's application in advertising coupled with participants' understandings of emotions as easily manipulated led to concerns that this manipulation could occur in marketing, giving entities the power to more strongly influence people's purchasing behavior.

Individual risk: some are prone to harm more than others. Participants noted how some individuals can experience harm more than others, leading to *unfair* outcomes. For instance, P7 discussed how some individuals would be harmed more than others in emotion recognition-enabled marketing: "*I mean, I suppose if they were using the data to purposely advertise expensive stuff to people that were feeling super vulnerable. I feel like that's harmful*." On the notion of how different people may experience harm to different extents, P10 did not think there was a risk of them personally being harmed, but they could see how harm could come about, saying: "*I guess I couldn't think of any situations where I might experience harm. I know there are definitely people with far less privilege than I have, so they would more than likely definitely experience harm, whether it's an increase in their insurance rate or ...*" Participants noted that vulnerable social media users (e.g., those in vulnerable emotional states) can be disproportionately harmed, questioning emotion recognition's fairness on social media.

Individual risk: negative impacts on emotional and mental health. While some participants were generally concerned about social media use and content online affecting mental health, these concerns were exacerbated when thinking about emotion recognition. For example, P2 reflected on the newsfeed's impact on wellbeing, "*But mostly from the perspective of what shows up on my newsfeed…I've noticed in the past year that Facebook does have negative effects on my emotional health at times*." P2 further explained that she has depression and fears that emotion recognition might result in targeted content that exacerbates it: "*I'm afraid of the kinds of feedback loops that it could create and influence, not just my, but everybody's emotional health and the emotional state that they're in. I do suspect that, to some extent, this [emotion recognition to deliver content] is already going on*." Specifically related to emotions and emotion recognition, P2 elaborated: "*They're going to be seeing that post, and they're going to be seeing that I'm feeling depressed. They're going to be feeding into that because they're going to see that I'm drawn or attracted to articles that might make me feel even more depressed…Eventually there's this feedback loop where, 'Oh, this is a depressed person. We're gonna feed this stuff because they just seem to gobble up this information*.'" First, this example illustrates how some participants believed that the kinds of content they receive is because of the kinds of content they consume, for example in the mental health context. Second, emotion recognition-enabled content delivery on social media can be *particularly* harmful to those experiencing mental health challenges.

Others noted how temporal aspects of people's emotions can further complicate emotion recognition and its impacts when used to deliver content. As P6 said: "*Because a person can be in a different mindset or in a different space or they may be better than they were from what they were a year ago or two years ago or they could be worse off. So if you're trying to market based on, say for instance, this person was in a bad place a year ago and you're trying to market something that has to do with overcoming drug addiction, rape, or anything, and they've forgotten about that and now you want to market them something about overcoming heroin addiction or something, that could take a person into a negative head space where they're now into a positive head space*." People's mental and emotional states change all the time. When and how (if at all) emotion recognition should be used to deliver mental health-related content is not a given and is an important area for further research.

Individual risk: identity and digital image misrepresentation across time. Participants raised concerns about the image that emotion recognition on social media will create about them online, including representations that will live beyond their lifetime. The lack of control over what persona is created about them and their emotional states and reactions, especially when skewed in some way, was one dimension of this risk. For example, P10 said: "*I do think that is problematic, especially if it becomes an emotion of record, like it's somewhere on the system that the machine predicted I felt this way, but I expressed nothing, and that's the only way anyone can know how I felt in that way, like it's the year 2200, I'm not here anymore, people all assume I felt a particular way*." Similarly, P6 believed that emotion recognition-enabled personas would not fairly represent them and their emotions: "*…that can make me to be somebody that I'm completely not*."

Participants were also concerned about misinterpreted data and its impacts. As P1 said: "*I can see that becoming a slippery slope of misinterpreted data…So much of the stuff we think is done by computers and done really well, is either poorly done or done by people. I think the trying to rely on that to make decisions in the future to learn about people is kind of playing with fire I think. There are a lot of mistakes that can happen really easily*." P5 echoed similar concerns around the harm that false positives could cause: "*I guess there is the potential to use that to identify people that could be risks in society, but at the same respect…there is also a problem in looking for things that haven't occurred yet. Even though there's the potential for something, it might never occur*." The possibility for unfair and inaccurate interpretations and lack of control over one's digitally curated image as enabled by emotion recognition can impact people during and well beyond their life time.

Individual risk: challenges with holding algorithms responsible. Participants raised concerns about a perceived lack of responsibility and regulation with algorithms employed by social media companies. For instance, P3 said: "*In the end it [the algorithm] can tell you to do things, but the thing is if it tells me to do something and I do it, and completely fucks me over, there is no accountability. Just all of mine*." P3 further elaborated that: "*It's different if a person tells you to do something, there's laws against that…*" Participants understood risks around emotion recognition in part by comparing the existing legal and policy infrastructure within which they exist with other legal structures that are more familiar to them (e.g., in-person civilian disputes). While P3 raised these concerns, they were not against algorithms categorically, but wanted them to be used responsibly, noting: "*Social media, it's everyone's journals or whatever, their lives are online in a way and we should handle them just as carefully as if you're actually sitting down with someone trying to talk to them*." How to make algorithms accountable is an ongoing debate in Computing (e.g., Human-Computer Interaction, Social Computing, Fairness, Accountability, and Transparency). We highlight a need for this debate in the emotion recognition context.

Societal risk. Beyond concern for the individual, some were concerned about what emotion recognition could mean for others or society as a whole on a broader level. For example, P8 said: *"I think it would make me wonder about the state of the world and the state of my world, to have that, to have my emotions manipulated*." P2 elaborated on detrimental impacts emotion recognition can have on democracy because of companies' political powers: "*The social level of using that information for political or social control, that bothers me a lot…I think having large companies that can in essence understand what we're feeling, and manipulate what we're feeling, provides greater and greater potential for a fascist or totalitarian regime to build in the country. Or for political unrest to be provoked by using these kinds of means. These companies are huge and they already have a lot of political power. Not all of the people that have the most power in these companies are good people*." This account may resonate with what was learned in the Cambridge Analytica case that used emotional and psychological profiling based on social media data to deliver content and sway political opinions [55]. These comments highlight the notion that emotion recognition's impact can go beyond the individual and can have political and social impacts.

DISCUSSION

We apply a Science and Technology Studies (STS) lens to emotion recognition to situate our work within the broader sociotechnical scholarship. We argue that it is crucial to account for the *humans* whose emotions these technologies recognize in a timely manner. We discuss implications for algorithmic accountability and advocate for considering emotion data as sensitive in research and practice.

Emotion recognition as *socio*technical. "Cultural lag" [91] refers to the fast growth in technology and slower speed of developing guidelines of ethical use. Failing to develop social consensus on ethical uses of emerging technologies leads to breakdown in social solidarity and rise of social conflict, and impact privacy rights [80]. We argue that emotion recognition is one context in which we need to avoid cultural lag. It is a kind of algorithmic decision-making, and is expected to reach its peak between 2021 and 2024 [36]. In 2019, considering the state of emotion recognition technologies, we have little time left in what Social Constructivism [95] calls the *interpretive flexibility* period. In this period, different interpretations of the emerging technology emerge from *relevant social groups* (groups with opinions about what problems technology should address). This period is followed by *stabilization* (when several technologies are developed to address the problem) and *closure* (when the relevant social group considers the problem solved). It is *before* the stabilization period, and certainly before the closure period, that we should critically decide what problems are important and what social groups should be included in decisions about developing, making sense of, and adopting emerging technologies [105]. People may have concerns about emerging technologies. While they may acclimate to new technology, technology could also shift to meet their demands. Such negotiation is arguably easier in earlier phases (e.g., interpretive flexibility). A first step is understanding people's concerns.

In this work, we began to uncover the concerns and perspectives of an important relevant social group: the *humans* who provide the data that make emotion recognition technologies possible, and who can be impacted by this technology in profound ways. It is important to account for this relevant social group's needs rather than *assuming* what constitutes their welfare [45]. Accounting for the social context in designing technologies is important as it leads to more *fair* systems, and we should avoid *closure* until we have addressed concerns from a diversity of social groups [105]. We account for the social context of emotion recognition technologies in this work. We contribute novel understandings of social media users' attitudes and perceived risks and outcomes in relation to emotion recognition on social media. The majority of participants were uncomfortable with emotion recognition, and this discomfort was often related to concerns over privacy, consent, agency, and potential harm. Our goal here was not to chart how emotion recognition technology should be designed, but rather, to identify people's attitudes towards it in this point of time when we have not yet reached *closure*. Our findings provide empirical evidence that technologists and academics building and designing emotion recognition, as well as policy makers, can refer to if they aspire to foreground humans' values and concerns in their work.

Emotion recognition and algorithmic accountability. Algorithmic accountability is about assigning responsibility for how algorithms are created, their societal impacts, and potential harms [3,23,52,110]. Therefore, we essentially identified algorithmic accountability's dimensions for emotion recognition related to its impacts and harms. Other emerging AI technologies subject to scholarly critique for their impact and harm include: facial emotion recognition (due to bias [99,114,115] e.g., showing no matter how much a Black person smiles, they are identified with more negative emotions [99]) and Automatic Gender Recognition (due to compromising privacy and autonomy [59]). We extend these works to the emotion recognition context by uncovering risks (including and beyond privacy and autonomy), and flagging how some may be harmed more than others.

Specifically, participants noted a lack of shared responsibility from algorithms and social media companies, which informed their attitudes towards emotion recognition on social media and signified risks. Moreover, our analysis identified being manipulated, unfairness in harm distribution (i.e., some would be more prone to harm than others), negative impacts on mental health (especially for those already experiencing mental health challenges), and losing autonomy over how one is digitally represented over time as anticipated risks associated with emotion recognition on social media. Human autonomy is the ability "to be one's own person" and impacts wellbeing [31]. A possible outcome we uncovered was leaving social media or limiting use. On a high level, this anticipated outcome parallels Foucault's notion of the 'panopticon' where people do not know if they are being watched and behave as though they are [51]. Surveillance threats have chilling effects on online participation [50]. More specifically, this outcome can harm those who rely on social media for social support and community, which impacts wellbeing [58] – especially for marginalized individuals who may not find support elsewhere. Feeling safe to be and express one's self and emotions also improves wellbeing [93]. A "harm-reduction framework" for algorithmic fairness argues that algorithm's effects on individuals' wellbeing should be considered [3]. In this sense, our findings shed light on what harms emotion recognition technologies should account for to be fair. By turning our focus on to the *humans* in emotion recognition, we encourage technologists in this space to move towards a stronger emphasis on the humans involved, with all their complexity, and attend to *their* wellbeing and concerns.

Emotional data as sensitive. A recent Pew survey [109] found that users' comfort level with social media companies using their data depends on *what* it is used for, covering four contexts but not specifying the emotional nature of the data. Our findings show that when considering *emotional* data, it is only *sometimes* (not always) that people's attitudes depend on what their data is used for. Additionally, privacy research identifies data related to health, location, web browsing, age, finances, and private communication to be highly sensitive [17,77,78,123]. While data in these contexts can have emotional dimensions to them, data can also be about one's emotions within and beyond these contexts. Our findings show that data about emotions, and data with emotional implications, are also highly sensitive and vulnerable. This has implications for researchers, technologists, and policy makers alike in deciding what data to treat as sensitive. We argue that emotional data is a type of data that warrants particular and explicit attention in research and practice.

CONCLUSION

We examined social media users' attitudes towards emotion recognition to contribute to our knowledge about socially and ethically responsible use and treatment of data in algorithmic advancements that impact humans' personal lives. We uncovered the ways people conceived of their emotions and emotion data when considering them being harvested by algorithms, and how these conceptions inform attitudes towards emotion recognition on social media. We identified outcomes and risks associated with emotion recognition on social media as perceived by participants, highlighting their values towards emotion recognition on social media. We argue that technologies that see into and infer insight from and about people's most vulnerable moments and emotional, private data should acknowledge humans' complexities. Acknowledging this complexity, does not necessarily mean building more accurate technologies to infer those complexities; rather, these innovations must prioritize the preferences, desires, and values of the *humans* they impact.

REFERENCES

1. Puneet Agrawal. 2017. Emotionally connected responses from a digital assistant. Retrieved June 26, 2019 from https://patents.google.com/patent/WO2017078960A1/en
2. Hessa AlBalooshi, Shahram Rahmanian, and Rahul Venkatesh Kumar. 2018. EmotionX-SmartDubai_NLP: Detecting User Emotions In Social Media Text. In *Proceedings of the Sixth International Workshop on Natural Language Processing for Social Media*, 45–49. Retrieved June 17, 2019 from https://www.aclweb.org/anthology/W18-3508
3. Micah Altman, Alexandra Wood, and Effy Vayena. 2018. A Harm-Reduction Framework for Algorithmic Fairness. *IEEE Security & Privacy* 16, 3: 34–45. https://doi.org/10.1109/MSP.2018.2701149
4. Tawfiq Ammari, Jofish Kaye, Janice Y. Tsai, and Frank Bentley. 2019. Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Transactions on Computer-Human Interaction* 26, 3: 1–28. https://doi.org/10.1145/3311956
5. Tawfiq Ammari, Meredith Ringel Morris, and Sarita Yardi Schoenebeck. 2014. Accessing Social Support and Overcoming Judgment on Social Media among Parents of Children with Special Needs. In *Proceedings of ICWSM 2014*.
6. Nazanin Andalibi. What Happens After Disclosing Stigmatized Experiences on Identified Social Media: Individual, Dyadic, and Social/Network Consequences. In *Under Review for CHI 2019*.
7. Nazanin Andalibi and Andrea Forte. 2018. Announcing Pregnancy Loss on Facebook: A Decision-Making Framework for Stigmatized Disclosures on Identified Social Network Sites. In *CHI 2018*. https://doi.org/10.1145/3173574.3173732
8. Nazanin Andalibi and Andrea Forte. 2018. Responding to Sensitive Disclosures on Social Media: A Decision-Making Framework. *ACM Transactions on Computer-Human Interaction (TOCHI)*. http://dx.doi.org/10.1145/3241044
9. Nazanin Andalibi, Margaret E. Morris, and Andrea Forte. 2018. Testing Waters, Sending Clues: Indirect Disclosures of Socially Stigmatized Experiences on Social Media. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW: 1–23. https://doi.org/10.1145/3274288
10. Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive Self-disclosures, Responses, and Social Support on Instagram: The Case of #Depression. In *Proceedings of the Conference on Computer Supported Cooperative Work & Social Computing* (CSCW ’17). https://doi.org/10.1145/2998181.2998243
11. Ian Barnett and John Torous. 2019. Ethics, Transparency, and Public Health at the Intersection of Innovation and Facebook’s Suicide Prevention Efforts. *Annals of Internal Medicine*. https://doi.org/10.7326/M19-0366
12. Solon Barocas and Danah Boyd. 2017. Engaging the Ethics of Data Science in Practice. *Commun. ACM* 60, 11: 23–25. https://doi.org/10.1145/3144172
13. Lisa Feldman Barrett. 2006. Are Emotions Natural Kinds? *Perspectives on Psychological Science* 1, 1: 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x
14. Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella, Aleix M. Martinez, and Seth D. Pollak. 2019. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements: *Psychological Science in the Public Interest*. https://doi.org/10.1177/1529100619832930
15. Eric P.S. Baumer, Timothy Berrill, Sarah C. Botwinick, Jonathan L. Gonzales, Kevin Ho, Allison Kundrik, Luke Kwon, Tim LaRowe, Chanh P. Nguyen, Fredy Ramirez, Peter Schaedler, William Ulrich, Amber Wallace, Yuchen Wan, and Benjamin Weinfeld. 2018. What Would You Do?: Design Fiction and Ethics. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork* (GROUP ’18), 244–256. https://doi.org/10.1145/3148330.3149405
16. Natalya N. Bazarova, Yoon Hyung Choi, Victoria Schwanda Sosik, Dan Cosley, and Janis Whitlock. 2015. Social sharing of emotions on Facebook: Channel differences, satisfaction, and replies. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 154–164. Retrieved from http://dl.acm.org/citation.cfm?id=2675297
17. Igor Bilogrevic, Kévin Huguenin, Berker Agir, Murtuza Jadliwala, Maria Gazaki, and Jean-Pierre Hubaux. 2016. A machine-learning based approach to privacy-aware information-sharing in mobile social networks. *Pervasive and Mobile Computing* 25: 125–142. https://doi.org/10.1016/j.pmcj.2015.01.006
18. Robert Brannon. 1976. The male sex role — And what it’s done for us lately. In *The forty-nine percent majority: The male sex role* (Edited by Deborah S. David and Robert Brannon). Addison-Wesley, Reading, MA.
19. John S. Breese and John Eugene Ball. 1999. Modeling a user’s emotion and personality in a computer user interface. Retrieved June 26, 2019 from https://patents.google.com/patent/US5987415A/en
20. Jed R. Brubaker, Lynn S. Dombrowski, Anita M. Gilbert, Nafiri Kusumakaulika, and Gillian R. Hayes. 2014. Stewarding a legacy: responsibilities and relationships in the management of post-mortem data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Pages 4157-4166*, 4157–4166. https://doi.org/10.1145/2556288.2557059
21. Manuel G. Calvo and Lauri Nummenmaa. 2016. Perceptual and affective mechanisms in facial expression recognition: An integrative review.

Cognition & Emotion 30, 6: 1081–1106. https://doi.org/10.1080/02699931.2015.1049124

22. R. A. Calvo and S. D'Mello. 2010. Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing* 1, 1: 18–37. https://doi.org/10.1109/T-AFFC.2010.1
23. Robyn Caplan, Lauren Hanson, Joan Donovan, and Jeanna Matthews. 2018. *Tech Algorithm Briefing: How Algorithms Perpetuate Racial Bias and Inequality*. Data & Society. Retrieved from https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL-4.pdf
24. J. M. Carroll. 2000. Five reasons for scenario-based design. *Interacting with Computers* 13, 1: 43–60. https://doi.org/10.1016/S0953-5438(00)00023-0
25. Stevie Chancellor, Eric P.S. Baumer, and Munmun De Choudhury. 2019. Who is the "Human" in Human-Centered Machine Learning: The Case of Predicting Mental Health from Social Media. *Proc. ACM Hum.-Comput. Interact*. https://doi.org/10.1145/3359249
26. Sheldon Chang. 2018. Determining a mood for a group. Retrieved August 7, 2019 from https://patents.google.com/patent/US10061977/en
27. Hsuan-Ting Chen. 2018. Revisiting the Privacy Paradox on Social Media With an Extended Privacy Calculus Model: The Effect of Privacy Concerns, Privacy Self-Efficacy, and Social Capital on Privacy Management. *American Behavioral Scientist* 62, 10: 1392–1412. https://doi.org/10.1177/0002764218792691
28. Munmun De Choudhury, Scott Counts, and Michael Gamon. 2012. Not All Moods Are Created Equal! Exploring Human Emotional States in Social Media. In *ICWSM*.
29. Munmun De Choudhury, Michael Gamon, and Scott Counts. 2012. Happy, Nervous or Surprised? Classification of Human Affective States in Social Media. In *Sixth International AAAI Conference on Weblogs and Social Media*. Retrieved July 15, 2019 from https://www.microsoft.com/en-us/research/publication/happy-nervous-or-surprised-classification-of-human-affective-states-in-social-media/
30. Sven-Ake Christianson (ed.). 1992. *The Handbook of Emotion and Memory: Research and Theory*. Psychology Press, Hillsdale, N.J.
31. John Christman. 2018. Autonomy in Moral and Political Philosophy. In *The Stanford Encyclopedia of Philosophy* (Spring 2018), Edward N. Zalta (ed.). Metaphysics Research Lab, Stanford University. Retrieved August 7, 2019 from https://plato.stanford.edu/archives/spr2018/entries/autonomy-moral/
32. Mike Conway. 2014. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. *Journal of Medical Internet Research* 16, 12: e290. https://doi.org/10.2196/jmir.3617
33. Mike Conway and Daniel O'Connor. 2016. Social Media, Big Data, and Mental Health: Current Advances and Ethical Implications. *Current Opinion in Psychology* 9: 77–82. https://doi.org/10.1016/j.copsyc.2016.01.004
34. Kate Crawford and Jason Schultz. 2013. *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*. Social Science Research Network, Rochester, NY. Retrieved August 7, 2019 from https://papers.ssrn.com/abstract=2325784
35. Baobao Zhang and Allan Dafoe and Center for the Governance of AI Oxford Future of Humanity Institute, University of. 2019. *Artificial Intelligence: American Attitudes and Trends*. Retrieved June 13, 2019 from https://governanceai.github.io/US-Public-Opinion-Report-Jan-2019/
36. Melissa Davis, Gareth Herschel, and Jim Hare. 2019. *Hype Cycle for Customer Experience Analytics, 2019*. Gartner. Retrieved August 8, 2019 from https://www.gartner.com/document/3956037?ref=solrAll&refval=227857318&qid=7df2169fe08cf98b6fda16
37. Munmun De Choudhury, Scott Counts, Eric J Horvitz, and Aaron Hoff. 2014. Characterizing and predicting postpartum depression from shared facebook data. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing*, 626–638.
38. Tony Dokoupil. 2018. Researcher who inspired Cambridge Analytica's data harvesting says era of privacy is over. Retrieved August 7, 2019 from https://www.cbsnews.com/news/michal-kosinski-researcher-who-inspired-facebook-data-harvesting-says-era-of-privacy-is-over/
39. Dana S Dunn, Kristen R Clinton Goldbach, Judith N Lasker, and Lori J Toedter. 1991. Explaining pregnancy loss: Parents' and physicians' attributions. *OMEGA-Journal of Death and Dying* 23, 1: 13–23.
40. J. A. Easterbrook. 1959. The effect of emotion on cue utilization and the organization of behavior. *Psychological Review* 66, 3: 183–201. https://doi.org/10.1037/h0047707
41. Paul Ekman. 2004. Emotions revealed. *BMJ* 328, Suppl S5: 0405184. https://doi.org/10.1136/sbmj.0405184
42. Paul Ekman and Wallace V Friesen. 1975. Unmasking the face: A guide to recognizing emotions from facial cues.
43. Nicole B. Ellison, Charles Steinfield, and Cliff Lampe. 2007. The Benefits of Facebook "Friends:" Social Capital and College Students' Use of Online Social Network Sites. *Journal of Computer-Mediated Communication* 12, 4: 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

44. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. "I Always Assumed That I Wasn'T Really That Close to [Her]": Reasoning About Invisible Algorithms in News Feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), 153–162. https://doi.org/10.1145/2702123.2702556
45. Virginia Eubanks. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
46. Casey Fiesler and Blake Hallinan. 2018. "We Are the Product": Public Reactions to Online Data Sharing and Privacy Controversies in the Media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–13. https://doi.org/10.1145/3173574.3173627
47. Casey Fiesler and Nicholas Proferes. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media + Society* 4, 1: 2056305118763366. https://doi.org/10.1177/2056305118763366
48. Janet Finch. 1987. The Vignette Technique in Survey Research. *Sociology* 21, 1: 105–114. https://doi.org/10.1177/0038038587021001008
49. Brett Q. Ford and James J. Gross. 2019. Why Beliefs About Emotion Matter: An Emotion-Regulation Perspective. *Current Directions in Psychological Science* 28, 1: 74–81. https://doi.org/10.1177/0963721418806697
50. Andrea Forte, Nazanin Andalibi, and Rachel Greenstadt. 2017. Privacy, Anonymity, and Perceived Risk in Open Collaboration: A Study of Tor Users and Wikipedians. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 1800–1811. https://doi.org/10.1145/2998181.2998273
51. Michel Foucault. 2012. *Discipline and Punish: The Birth of the Prison*. Knopf Doubleday Publishing Group.
52. Simson Garfinkel, Jeanna Matthews, Stuart S Shapiro, and Jonathan M Smith. 2017. Toward Algorithmic Transparency and Accountability. *Communications of the ACM 60*. Retrieved August 6, 2019 from https://cacm.acm.org/magazines/2017/9/220423-toward-algorithmic-transparency-and-accountability/fulltext
53. Maria Gendron and Lisa Feldman Barrett. 2009. Reconstructing the Past: A Century of Ideas About Emotion in Psychology. *Emotion Review* 1, 4: 316–339. https://doi.org/10.1177/1754073909338877
54. Ysabel Gerrard and Tarleton Gillespie. 2019. When Algorithms Think You Want to Die. *Wired*. Retrieved February 21, 2019 from https://www.wired.com/story/when-algorithms-think-you-want-to-die/
55. Kevin Granville. 2018. Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens - The New York Times. Retrieved from https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html
56. Michael Froimowitz Greenzeiger, Ravindra Phulari, and Mehul K. Sanghavi. 2015. Inferring user mood based on user and group characteristic data. Retrieved June 26, 2019 from https://patents.google.com/patent/US8965828B2/en
57. Oliver L. Haimson. 2019. Mapping gender transition sentiment patterns via social media data: toward decreasing transgender mental health disparities. *Journal of the American Medical Informatics Association*. https://doi.org/10.1093/jamia/ocz056
58. Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2015. Disclosure, Stress, and Support During Gender Transition on Facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '15), 1176–1190. https://doi.org/10.1145/2675133.2675152
59. Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M. Branham. 2018. Gender Recognition or Gender Reductionism?: The Social Implications of Embedded Gender Recognition Systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–13. https://doi.org/10.1145/3173574.3173582
60. Zenobia Hedge. 2018. Emotion recognition and sentiment analysis market to reach $3.8bn by 2025, says Tractica. *IoT Now - How to run an IoT enabled business*. Retrieved August 7, 2019 from https://www.iot-now.com/2018/03/08/78263-emotion-recognition-sentiment-analysis-market-reach-3-8bn-2025-says-tractica/
61. David M. Herszenhorn. 2014. Heightened Security, Visible and Invisible, Blankets the Olympics. *The New York Times*. Retrieved August 7, 2019 from https://www.nytimes.com/2014/02/14/sports/olympics/heightened-security-visible-and-invisible-blankets-the-olympics.html
62. Andrew C. High, Anne Oeldorf-Hirsch, and Saraswathi Bellur. 2014. Misery rarely gets company: The influence of emotional bandwidth on supportive communication on Facebook. *Computers in Human Behavior* 34, 0: 79–88. https://doi.org/10.1016/j.chb.2014.01.037
63. Rhidian Hughes. 1998. Considering the Vignette Technique and its Application to a Study of Drug Injecting and HIV Risk and Safer Behaviour. *Sociology of Health & Illness* 20, 3: 381–400. https://doi.org/10.1111/1467-9566.00107
64. Molly Jackman and Lauri Kanerva. 2016. Evolving the IRB: Building Robust Review for Industry Research. *Washington and Lee Law Review Online* 72, 3: 442.

65. Rana el Kaliouby and Daniel Abraham Bender. 2014. Mental state well being monitoring. Retrieved August 6, 2019 from https://patents.google.com/patent/US20140200463A1/en?assignee=affectiva&page=2
66. Rana el Kaliouby, Rosalind Picard, and Simon Baron-Cohen. 2006. Affective computing and autism. *Annals of the New York Academy of Sciences* 1093: 228–248. https://doi.org/10.1196/annals.1382.016
67. Youssef Kashef, Rana El Kaliouby, Ahmed Adel OSMAN, Niels Haering, and Viprali Bhatkar. 2014. Mental state analysis using heart rate collection based video imagery. Retrieved August 6, 2019 from https://patents.google.com/patent/WO2014145204A1/en
68. Yaser Khan, Chris HUYBREGTS, Jaeyoun Kim, and Thomas C. Butcher. 2016. Real-time emotion recognition from audio signals. Retrieved June 11, 2019 from https://patents.google.com/patent/WO2016014321A1/en
69. Y. Kim, H. Lee, and E. M. Provost. 2013. Deep learning for robust feature generation in audiovisual emotion recognition. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 3687–3691. https://doi.org/10.1109/ICASSP.2013.6638346
70. Peter Kindness, Judith Masthoff, and Chris Mellish. 2017. Designing emotional support messages tailored to stressors. *International Journal of Human-Computer Studies* 97: 1–22. https://doi.org/10.1016/j.ijhcs.2016.07.010
71. Funda Kivran-Swaine, Jeremy Ting, Jed R. Brubaker, Rannie Teodoro, and Mor Naaman. 2014. Understanding Loneliness in Social Awareness Streams: Expressions and Responses. In *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*.
72. Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24: 8788–8790. https://doi.org/10.1073/pnas.1320040111
73. Akash Krishnan and Matthew Fernandez. 2013. System and method for recognizing emotional state from a speech signal. Retrieved June 26, 2019 from https://patents.google.com/patent/US8595005B2/en
74. Joseph Lindley and Paul Coulton. 2015. Back to the Future: 10 Years of Design Fiction. In *Proceedings of the 2015 British HCI Conference* (British HCI ’15), 210–211. https://doi.org/10.1145/2783446.2783592
75. Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone. 2008. Physiology-based affect recognition for computer-assisted intervention of children with Autism Spectrum Disorder. *International Journal of Human-Computer Studies* 66, 9: 662–677. https://doi.org/10.1016/j.ijhsc.2008.04.003
76. George Loewenstein and Jennifer S. Lerner. 2003. The role of affect in decision making. In *Handbook of affective sciences*. Oxford University Press, New York, NY, US, 619–642.
77. Mary Madden. 2014. Americans Consider Certain Kinds of Data to be More Sensitive than Others | Pew Research Center. Retrieved August 7, 2019 from https://www.pewinternet.org/2014/11/12/americans-consider-certain-kinds-of-data-to-be-more-sensitive-than-others/
78. Miguel Malheiros, Sören Preibusch, and M. Angela Sasse. 2013. “Fairly Truthful”: The Impact of Perceived Effort, Fairness, Relevance, and Sensitivity on Personal Data Disclosure. In *Trust and Trustworthy Computing* (Lecture Notes in Computer Science), 250–266.
79. Lydia Manikonda and Munmun De Choudhury. 2017. Modeling and Understanding Visual Attributes of Mental Health Disclosures in Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI ’17), 170–181. https://doi.org/10.1145/3025453.3025932
80. Kimball P. Marshall. 1999. Has Technology Introduced New Ethical Problems? *Journal of Business Ethics* 19, 1: 81–90. https://doi.org/10.1023/A:1006154023743
81. Michael James Matas, Michael Waldman Reckhow, and Yaniv Taigman. 2017. Systems and methods for dynamically generating emojis based on image analysis of facial features. Retrieved June 26, 2019 from https://patents.google.com/patent/US20170140214A1/en
82. Louise Matsakis. 2018. Online Ad Targeting Does Work—As Long As It’s Not Creepy. *Wired*. Retrieved February 22, 2019 from https://www.wired.com/story/online-ad-targeting-does-work-as-long-as-its-not-creepy/
83. Adam McCann. 2019. Most Diverse Cities in the U.S. *WalletHub*. Retrieved August 3, 2019 from https://wallethub.com/edu/most-diverse-cities/12690/
84. Andrew McStay. 2018. *Emotional AI: The Rise of Empathic Media*. SAGE.
85. Jacob Metcalf, Emily F. Keller, and danah boyd. 2019. *Perspectives on Big Data, Ethics, and Society*. Council for Big Data, Ethics, and Society. Retrieved August 7, 2019 from https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/
86. Michelle N. Meyer. 2015. *Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation*. Social Science Research Network, Rochester, NY. Retrieved August 7, 2019 from https://papers.ssrn.com/abstract=2605132

87. Jude Mikal, Samantha Hurst, and Mike Conway. 2016. Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. *BMC medical ethics* 17: 22. https://doi.org/10.1186/s12910-016-0105-5
88. Dan Muriello, Lizzy Donahue, Danny Ben-David, Umut Ozertem, and Reshef Shilon. 2018. Under the hood: Suicide prevention tools powered by AI. *Facebook Code*. Retrieved January 29, 2019 from https://code.fb.com/ml-applications/under-the-hood-suicide-prevention-tools-powered-by-ai/
89. Anthony Nadler and Lee McGuigan. 2018. An impulse to exploit: the behavioral turn in data-driven marketing. *Critical Studies in Media Communication* 35, 2: 151–165. https://doi.org/10.1080/15295036.2017.1387279
90. Barak Reuven Naveh. 2015. Techniques for emotion detection and content delivery. Retrieved from https://patents.google.com/patent/US20150242679A1/en
91. William Fielding Ogburn. 1922. *Social Change with Respect to Culture and Original Nature*. B.W. Huebsch, Incorporated.
92. S. Panwar, P. Rad, K.-K.R. Choo, and M. Roopaei. 2019. Are you emotional or depressed? Learning about your emotional state from your music using machine learning. *Journal of Supercomputing* 75, 6: 2986–3009. https://doi.org/10.1007/s11227-018-2499-y
93. James W Pennebaker. 1995. *Emotion, disclosure, & health*. American Psychological Association.
94. Elizabeth A. Phelps, Sam Ling, and Marisa Carrasco. 2006. Emotion Facilitates Perception and Potentiates the Perceptual Benefits of Attention. *Psychological science* 17, 4: 292–299. https://doi.org/10.1111/j.1467-9280.2006.01701.x
95. Trevor J. Pinch and Wiebe E. Bijker. 1984. The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science* 14, 3: 399–441. https://doi.org/10.1177/030631284014003004
96. Nicholas Proferes. 2017. Information Flow Solipsism in an Exploratory Study of Beliefs About Twitter. *Social Media + Society* 3, 1: 2056305117698493. https://doi.org/10.1177/2056305117698493
97. Foster Provost and Tom Fawcett. 2013. Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data* 1, 1: 51–59. https://doi.org/10.1089/big.2013.1508
98. Andrew G. Reece and Christopher M. Danforth. 2017. Instagram photos reveal predictive markers of depression. *EPJ Data Science* 6, 1: 15. https://doi.org/10.1140/epjds/s13688-017-0110-z
99. Lauren Rhue. 2018. *Racial Influence on Automated Perceptions of Emotions*. Social Science Research Network, Rochester, NY. Retrieved March 13, 2019 from https://papers.ssrn.com/abstract=3281765
100. Bernard Rimé. 2009. Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review* 1, 1: 60–85.
101. James A. Russell. 1994. Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin* 115, 1: 102–141. https://doi.org/10.1037/0033-2909.115.1.102
102. Richard Scott Sadowsky, Rana el Kaliouby, Rosalind Wright Picard, Oliver Orion Wilder-Smith, Panu James Turcot, and Zhihong Zheng. 2011. Mental state analysis using web services. Retrieved August 6, 2019 from https://patents.google.com/patent/US20110301433A1/en?assignee=affectiva
103. Koustuv Saha, Larry Chan, Kaya De Barbaro, Gregory D. Abowd, and Munmun De Choudhury. 2017. Inferring Mood Instability on Social Media by Leveraging Ecological Momentary Assessments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3: 95:1–95:27. https://doi.org/10.1145/3130960
104. Irving Seidman. 2006. *Interviewing as qualitative research: a guide for researchers in education and the social sciences*. Teachers College Press, New York.
105. Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (FAT* ’19), 59–68. https://doi.org/10.1145/3287560.3287598
106. Irina Shklovski, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. 2014. Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (CHI ’14), 2347–2356. https://doi.org/10.1145/2556288.2557421
107. Natasha Singer. 2018. Creepy or Not? Your Privacy Concerns Probably Reflect Your Politics. *The New York Times*. Retrieved April 25, 2019 from https://www.nytimes.com/2018/04/30/technology/privacy-concerns-politics.html
108. Michael Skirpan and Casey Fiesler. 2018. Ad Empathy: A Design Fiction. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork* (GROUP ’18), 267–273. https://doi.org/10.1145/3148330.3149407
109. Aaron Smith. 2018. *Public Attitudes Toward Computer Algorithms | Pew Research Center*. Retrieved January 31, 2019 from http://www.pewinternet.org/2018/11/16/public-attitudes-toward-computer-algorithms/
110. Brad Smith and Harry Shum. 2018. *The Future Computed: Artificial Intelligence and its role in society*. Microsoft. Retrieved from https://blogs.microsoft.com/wp-content/uploads/2018/02/The-Future-Computed_2.8.18.pdf

111. Jay Stanleye. 2019. Experts Say “Emotion Recognition” Lacks Scientific Foundation. *American Civil Liberties Union*. Retrieved July 20, 2019 from https://www.aclu.org/blog/privacy-technology/surveillance-technologies/experts-say-emotion-recognition-lacks-scientific
112. Luke Stark. 2016. The emotional context of information privacy. *The Information Society* 32, 1: 14–27. https://doi.org/10.1080/01972243.2015.1107167
113. Luke Stark. 2018. Algorithmic psychometrics and the scalable subject. *Social Studies of Science* 48, 2: 204–231. https://doi.org/10.1177/0306312718772094
114. Luke Stark. 2018. Facial recognition, emotion and race in animated social media. *First Monday* 23, 9. https://doi.org/10.5210/fm.v23i9.9406
115. Luke Stark. 2019. Facial recognition is the plutonium of AI. *XRDS: Crossroads, The ACM Magazine for Students* 25, 3: 50–55. https://doi.org/10.1145/3313129
116. Anselm Strauss and Juliet Corbin. 1998. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. SAGE Publications.
117. J. Tang, Y. Zhang, J. Sun, J. Rao, W. Yu, Y. Chen, and A. C. M. Fong. 2012. Quantitative Study of Individual Emotional States in Social Networks. *IEEE Transactions on Affective Computing* 3, 2: 132–144. https://doi.org/10.1109/T-AFFC.2011.23
118. Tractica. Emotion Recognition and Sentiment Analysis | Tractica. Retrieved August 7, 2019 from https://www.tractica.com/research/emotion-recognition-and-sentiment-analysis/
119. Zeynep Tufekci. 2015. Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency. *Colorado Technology Law Journal* 13: 203.
120. Joseph Turow, Michael Hennessy, Nora Draper, and Ope Akanbi. 2018. *Divided We Feel: Partisan Politics Drive Americans’ Emotions Regarding Surveillance of Low-Income Populations*. Retrieved from https://repository.upenn.edu/asc_papers/543/
121. Blase Ur, Pedro Giovanni Leon, Lorrie Faith Cranor, Richard Shay, and Yang Wang. 2012. Smart, Useful, Scary, Creepy: Perceptions of Online Behavioral Advertising. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (SOUPS ’12), 4:1–4:15. https://doi.org/10.1145/2335356.2335362
122. Kristen Vaccaro, Dylan Huang, Motahhare Eslami, Christian Sandvig, Kevin Hamilton, and Karrie Karahalios. 2018. The Illusion of Control: Placebo Effects of Control Settings. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–13. Retrieved January 7, 2020 from http://dl.acm.org/doi/abs/10.1145/3173574.3173590
123. Yang Wang, Gregory Norice, and Lorrie Faith Cranor. 2011. Who Is Concerned about What? A Study of American, Chinese and Indian Users’ Privacy Concerns on Social Network Sites. In *Trust and Trustworthy Computing* (Lecture Notes in Computer Science), 146–153.
124. Yichen Wang and Aditya Pal. 2015. Detecting Emotions in Social Media: A Constrained Optimization Approach. In *IJCAI*.
125. Richmond Y. Wong 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), 121–133. https://doi.org/10.1145/2901790.2901801
126. Richmond Y. Wong, Deirdre K. Mulligan, and John Chuang. 2017. Using Science Fiction Texts to Surface User Reflections on Privacy. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers* (UbiComp ’17), 213–216. https://doi.org/10.1145/3123024.3123080
127. Richmond Y. Wong, Ellen Van Wyk, and James Pierce. 2017. Real-Fictional Entanglements: Using Science Fiction and Design Fiction to Interrogate Sensing Technologies. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (DIS ’17), 567–579. https://doi.org/10.1145/3064663.3064682
128. Meredith GF Worthen. 2013. An invitation to use craigslist ads to recruit respondents from stigmatized groups for qualitative interviews: *Qualitative Research*. https://doi.org/10.1177/1468794113481791
129. Michael Zimmer. 2010. “But the data is already public”: on the ethics of research in Facebook. *Ethics and Information Technology* 12, 4: 313–325. https://doi.org/10.1007/s10676-010-9227-5
130. Michael Zimmer. 2016. OkCupid Study Reveals the Perils of Big-Data Science. Retrieved August 7, 2019 from https://www.wired.com/2016/05/okcupid-study-reveals-perils-big-data-science/
131. Michael Zimmer. 2018. Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity. *Social Media + Society* 4, 2: 205630511876830. https://doi.org/10.1177/2056305118768300
132. Shoshana Zuboff. 2015. *Big Other: Surveillance Capitalism and the Prospects of an Information Civilization*. Social Science Research Network, Rochester, NY. Retrieved August 7, 2019 from https://papers.ssrn.com/abstract=2594754