Human-first, please: Assessing citizen views and industrial ambition for emotional AI in recommender systems

Bakir, Vian; Laffer, Alex; McStay, Andrew

Surveillance and Society

DOI:
10.24908/ss.v21i2.16015

Published: 03/07/2023

Publisher's PDF, also known as Version of record

Surveillance & Society

Article

Human-First, Please: Assessing Citizen Views and Industrial Ambition for Emotional AI in Recommender Systems

Vian Bakir
Bangor University, UK
v.bakir@bangor.ac.uk

Alexander Laffer
Bangor University, UK
a.laffer@bangor.ac.uk

Andrew McStay
Bangor University, UK
mcstay@bangor.ac.uk

Abstract

This paper qualitatively explores the views of diverse members of the British public on applications of biometric emotional AI technologies patented by two globally dominant consumer-facing recommender systems, Amazon and Spotify. Examining Amazon and Spotify patents for biometric profiling of users' emotions, disposition, and behaviour to offer them tailored services, ads, and products from their wider platforms, this paper points to industrial ambition regarding emotional AI. Little is known about ordinary people's views on deployment of such technology, and given the complex, abstract, and future-facing nature of such technologies, ascertaining informed lay views is hard. We address this through our innovative, qualitative study of diverse British-based adults (n=46) that presents to them near-horizon use cases in an interactive fictional narrative that deploys design fiction principles and ContraVision techniques. We find the themes of "usefulness," "resignation," "uneasy terms of engagement," and "human-first," adding rich and nuanced insights to prior survey work on users' views towards biometric-based emotional AI technologies. In contributing to a richer understanding of whether emotional AI technologies should be deployed in consumer-facing recommender systems, and if so, on what terms, we find that well-established policy-friendly criticisms apply to global emotional AI recommender systems. We conclude, however, that problems of alienation and need for a human-first approach to emerging AI technology are the most significant criticisms.

Introduction

Recommender systems simplify discovery and selection of items by using a ranking function to present a small subset of available items. Seemingly benign in helping consumers negotiate information overload, every online user has interacted with a recommender system, often without full awareness that their choices (for instance, on items bought or songs listened to) were invisibly nudged by recommendation engines. Increasingly, such recommender systems have signalled (through patents) their intention to incorporate users' datafied emotions by deploying emotional Artificial Intelligence (emotional AI) to improve their personalisation and targeting systems. As yet, little is known about these industrial ambitions or about people's views on deployment of such intimate technology, these forming the subject of this paper.

"Emotional AI" involves machine learning technologies, employed to label and react to human emotions. They do so via sensing words and images (such as sentiment analysis) and via sensing bodily behaviours (biometrics) including voice, facial expressions, gaze direction, gestures, heart rate, body temperature, respiration, and dermal electrical properties (McStay 2018). Only a few years ago, practical use cases of biometric-based emotional AI were rare, but this is changing in consumer-facing sectors such as cars (rolled out by legacy car companies to improve cabin experience and safety) and wearables (deployed by market

Bakir, Vian, Alexander Laffer, and Andrew McStay. 2023. Human-First, Please: Assessing Citizen Views and Industrial Ambition for Emotional AI in Recommender Systems. *Surveillance & Society* 21 (2): 205-222.
https://ojs.library.queensu.ca/index.php/surveillance-and-society/index | ISSN: 1477-7487

leaders such as Garmin to help users manage their mental health and day) (McStay 2018, 2023). Deployment of biometric-based emotional AI is also envisaged in the patents of global digital platforms such as Amazon (the world's largest online retailer and marketplace) and Spotify (the world's largest music streaming service provider). Their recommendation engines will be familiar to users of these services, notably Amazon's "Customers Who Bought This Item Also Bought…" feature, and Spotify's "Discover" button. As we will show, patents logged by Amazon and Spotify envisage using the emotion recognition of users to then offer them highly tailored services, ads, and products from their wider platforms.

The policy and governance context to recommender systems that function by emotional AI is also noteworthy. Emotional AI is of increasing concern to European Union regulators, as seen in the European Union's draft AI Act (European Commission 2021). In its current state of development, this risk-based legislation seeks to encourage innovation and trust in AI applications by applying variously lighter and stronger legal regimes to AI applications depending on their riskiness. It currently imposes specific transparency obligations on automated emotion recognition systems, seeing them as of limited risk if not used in high-risk settings (such as education, employment, justice, law, or immigration). These transparency obligations would apply to consumer-facing recommender systems deploying biometric emotion recognition. They entail telling users that the system is in use, to allow users "to make informed choices or step back from a given situation" (European Commission 2021: 14–15). Yet, little is known about the ambition of dominant platforms regarding deployment of such systems, or about people's views on any such deployment. Building on scholarship on emotional AI, capitalism, and user-based studies, our research questions are: (1) What does analysis of patents tell us about industrial ambition regarding emotional AI in dominant consumer-facing recommender systems? (2) Are people willing to engage with emotional AI in consumer-facing recommender systems, and if so, on what terms?

We answer this by analysing two industrial patents that advance emotional AI in globally significant recommender systems. We then deploy these insights in our innovative qualitative study of diverse British-based adults that enables us to consider their viewpoints on complex, abstract, future-facing technologies. To do so, the study develops an interactive narrative method using design fiction principles and ContraVision techniques to ground participants' experience with the proposed applications in everyday settings and help them think through the benefits and harms, while also minimising researcher bias.

We find that common across our diverse groups (purposively sampled based on age, ethnicity, and disability to surface diverse views) is the theme of "usefulness" (comprising positive and negative views). Where positive, this is policy-friendly, echoing policymakers' interest in choice and agency. Yet, other prominent themes are more negative including "human-first," "uneasy terms of engagement," and "resignation." We argue that it is the broader "human-first*"* insight that is most significant. It speaks to a more generalised sense of alienation, being out of control, and classic surveillance-based observations of technopoly and capitalism, albeit one mediated through biometrics and emotion. Throughout our study, we were sensitive to post-phenomenological criticism of the "general critique" of technology, with post-phenomenology advocating "grounded assessment" of specific technologies (Verbeek 2011). Nonetheless, our key finding *is* generalised concern expressed by diverse users, created by lack of control over specific aspects of system design and data processing.

We conclude that considering diverse users' views on emotional AI in recommender systems is vital in the development and governance of such systems. We also recommend that giving users a genuine choice in whether or not to use such systems is paramount. If people are resigned, or uneasy about the terms of engagement, then it cannot be sufficient to simply make users aware that an emotional AI system is in use and then leaving them to choose whether or not to use it (as currently proposed in the European Union's draft AI Act).

Emotional AI and Capitalism

The use of biometrics for authentication has become routine, and the body consequently and increasingly directly plays a role in commercial digital networked life. Adding emotion recognition systems to the mix of data streams routinely profiled in data-extractive forms of capitalism provides insights into users' decision-making processes (Barfar 2019). There are many terms for forms of capitalism presaged upon deploying data, metrics, and algorithmic management to extract value from immaterial labour and assets hitherto deemed unproductive or underproductive (Bakir and McStay 2018; Kienscherf 2022). For our purposes, Zuboff's (2015, 2019) popular term "surveillance capitalism" is useful in highlighting the behaviour modification (of users) that it enables. She states that surveillance capitalism has "exclusive capabilities to have access to behavioural surplus" (qtd. in Zuboff et al. 2019: 261) for the purposes of commodification, monetisation, and knowing, controlling, and modifying behavior; additionally, it constantly seeks new data-streams: "We need your heartbeat, your bloodstream,… your walk in the park,… how you sleep at night" (qtd. in Zuboff et al. 2019: 263). Exemplifying this, since 2014, Amazon has been "eavesmining" (a portmanteau of eavesdropping and datamining) to increase the functionality of its voice-activated personal assistant, Alexa (Neville 2020). West (2019) argues that Amazon seeks to make Alexa (and hence surveillance on its platform) an indispensable consumer service, thereby sweetening granular surveillance in once-private spaces and cultivating docile surveillant subjects. Indeed, to understand the speaker's intention, Amazon has also been exploring emotion-sensitive voice-based assistants based on prosody, namely variation in melody, intonation, pauses, stresses, intensity, vocal quality, and accents of speech (McStay 2018). Such comprehension of emotions would help Amazon interact more naturally and profile emotions in relation to users' requests while also priming the system to respond to users appropriately.

Zuboff's (2015; Zuboff et al. 2019) analysis, of course, is presaged by multiple surveillance scholars (Ball 2019). This includes von Otterlo (2014) on capitalists' attempt at massive scale behaviourism using interactive feedback loops of manipulation and behavioural engineering through data and algorithms. It includes Andrejevic (2011) on "affective economies," where marketers seek to manage consumers by datamining their sentiment alongside demographic and behavioural data. It includes Ball (2009) on the political economy of interiority, where institutions expose to others aspects of an individual's private world via datafication processes. With emotion data adding another potentially highly informative layer to extractive logics (Sadowski 2019) that see the primary resource of the big data industries as people themselves (Gregory and Sadowski 2021), we must also acknowledge the contribution of Autonomist Marxist scholarship, given its astuteness in recognising how subjectivity and emotional life is used to generate economic value. Berardi (2009: 109), for example, argues that "post-industrial life" is "marked by the submission of the soul, in which animated, creative, linguistic, emotional corporeality is subsumed and incorporated by the production of value" (see also Hardt and Negri 2000; Stiegler 2010). Others point to the role of capitalism in organising and modulating emotion and affective states for given goals, with Lazzarato (2014) seeing capitalist data processing in terms of "enslavement." While one can debate Lazzarato's (2014) word choice, the point that behaviour, affects, and subjectivity are rendered into machine-readable human capital and put to work at an industrial scale is plain (McStay 2018).

User-Based Studies

A growing amount of research investigates what people think of these extractive data arrangements. We discuss three areas of user-based studies covering (a) control over personal data, (b) recommender systems, and (c) emotional AI.

Repeatedly, studies find that users have issues controlling their personal data online. Hartman et al.'s (2020) British-based national survey finds that respondents prefer data governance approaches that give them control over their personal data, that include regulatory oversight, or that enable them to opt out of data gathering. Yet, other user-based studies find "privacy fatigue" (Hargittai and Marwick 2016) where people feel apathy or cynicism about online privacy because they believe that violations are inevitable and opting

out is impossible. Related, user-based studies in the US and UK also find "digital resignation" (Draper and Turow 2019) and "surveillance realism" (Dencik and Cable 2017) where people want to better control the information that digital entities hold about them but feel unable to do so given the surveillant power of corporations and governments and imbalances between citizen and institutional visibility. This fosters users' perception that there is no alternative system that could provide functionality without intrusive profiling.

Research into people's understanding of algorithms in society finds an "algorithm knowledge gap," even in digitally literate societies (Zarouali, Nelberger, and de Vreese 2021). However, studies on recommender system algorithms deployed on specific sites (such as Facebook's News Feed, digital news sites, and music recommender sites) find diverse understandings and practices (Eslami et al. 2016; Harambam et al. 2019; Shin 2020; Siles et al. 2020). For instance, studies from the Netherlands and South Korea find that people consider transparent, accurate systems for algorithmic curation on digital news sites convenient and useful, and they value possibilities to influence the recommendation algorithms (Harambam et al. 2019; Shin 2020). A Costa Rican user-based study finds that users variously see Spotify's recommender algorithms as internationally standardised or malleable towards more local cultural conceptions of friendship and social behaviour (Siles et al. 2020). Recent studies also document how users exert agency over algorithmic decision-making to advance their own interests, be this to craft an influencer online identity on Instagram (Cotter 2018); to develop musical identities and taste (Karakayali, Kostem, and Galip 2018); or, as heavy users, activists and artists, to submit to, resist, or disrupt algorithmic outputs (Siles et al. 2020; Velkova and Kaun 2021).

Regarding user-based studies on emotional AI, national demographically representative surveys on people's attitudes towards biometric-based emotional AI technologies conducted in the UK show a mix of results. Overall, a small majority would share their biometric and emotion-based data with consumer-facing sectors such as out-of-home outdoor advertisers and car manufacturers, but a small majority would not do so with organisations such as schools, workplaces, and political campaigners. Throughout, there is marked variation according to age: if care is taken over levels of personal identification, younger adults are more open to such technologies than older people (Bakir and McStay 2022). In another consumer-facing sector (toys), a survey that ascertains parental perspectives on networked toys that utilise data about the child's emotions finds ambivalence. Although more parents than not agree that a child-oriented wearable that tracks emotion is intrusive and express concern about what happens to the data, they also agree that it would help with parenting and want to be told on the box's packaging that it collects emotion data (McStay and Rosner 2021).

To summarise, user-based studies find that people want more control of their personal data online but do not expect to get this; have varied conceptions about, and practices towards, recommender algorithms; and hold a mix of attitudes towards emotional AI technologies (where negativity increases with age), with small majorities in favour of consumer-facing applications.

Materials and Methods

Analysing Patents

To analyse industrial ambition regarding emotional AI in consumer-facing recommender systems, we examine two patents from globally dominant platforms, Amazon and Spotify. While many other well-known companies log patents for emotion recognition applications, Amazon and Spotify were selected due to their scale of operation and interest in new forms of recommender systems. Patent analysis is a useful tool for understanding the ambition of secretive corporations engaged in surveillant capitalism (Iliadis and Acker 2022; McStay 2018). Patent filings afford critical scholars insight into the hopes, technical intentions, and worldviews of companies and owners, potentially unachievable by other means. Seemingly less guarded than other corporate public output, they not only give an upfront sense of what a company seeks to launch in the future but also its orientation and what it regards as acceptable. A patent is not a prediction, as a patent may be sought for many reasons, including intellectual property, scientific markers, ownership, and strategic

blocking of others innovating in similar technical domains. Yet, as a public statement, they say much about corporate values, soft technology ethics, and the types of technologies to emerge sooner or later. We describe key features of the patents deploying emotional AI logged by Amazon and Spotify below, noting what forms of biometric, emotional, and other contextual data they ingest and to what stated end. Significant features from these patents are then used in the fictional narrative that we present to our participants in workshops.

Patent by Amazon

In March 2017, Amazon filed a patent for "Voice-based determination of physical and emotional characteristics of users," divulging a potential use case for voice-controlled virtual assistants such as Amazon's Alexa (Jin and Wang 2018: 1). The patent observes that "physical conditions such as sore throats and coughs" and "emotional conditions such as an excited emotional state or a sad emotional state" may be determined partly on a user's voice input (Jin and Wang 2018: 7). The patent offers the example of the user telling Alexa that she is hungry. Alexa detects that the user's voice, which is coughing and sniffling, is different to usual, so asks if the user would like a chicken soup recipe and offers to order cough drops with one-hour delivery.

The patent asserts that the content offered by Alexa may be "highly targeted due to the real-time determination of the physical and/or emotional characteristics of the user, and may therefore be timely and relevant to the user's current state" (Jin and Wang 2018: 7). The various determinations made comprise the "meaning of user request," "background noise feature," and the user's "language accent," "physical status," and "emotional status." It notes that emotional states could comprise "happiness, joy, anger, sorrow, sadness, fear, disgust, boredom, stress, or other emotional states," and determinable health conditions could include "default or normal, sore throat, cold, thyroid issues, [or] sleepiness" (Jin and Wang 2018: 11). To enhance the targeting towards the device's user, it reveals that a user's physical and/or emotional characteristics would be combined with, "behavioral targeting criteria (e.g., browse history, number of clicks, purchase history, etc.) and/or contextual targeting criteria (e.g., keywords, page types, placement metadata, etc.)" (Jin and Wang 2018: 7). Of note, the patent explains that targeted content may include sponsored content and third-party content. In its provided example, if the user declines the chicken soup recipe, follow-up inquiries may be associated with particular advertisers. The patent suggests that Panera Bread (a US chain of bakery-café fast casual restaurants) "may be interested in presenting an audio advertisement to the user. As a result, the follow-up inquiry of 'would you like to order chicken soup?' may be directed towards determining whether the user is a target consumer for receiving a Panera Bread audio advertisement" (Jin and Wang 2018: 9).

Simulation of caring attention from Alexa for chicken soup in response to determining that the user has a cold, then, is simply a ruse to optimise and deliver a targeted ad or a suggestion to place a purchase order. Furthermore, the patent explains that content could be disseminated through all sorts of devices such as a television, laptop computer, tablet, computer monitor, speaker-type device, augmented reality, or virtual reality glasses (Jin and Wang 2018: 8), indicating that targeting may take place both online and offline—a highly pervasive surveillance scenario.

Patent by Spotify

A patent logged in 2021 by Spotify titled "Identification of Taste Attributes from an Audio Signal" underpins the second patent example that informs our user-based study. Spotify's goal is to improve its speech-enabled recommender services: namely, helping Spotify determine users' listening preferences to "provide personalized media recommendations" (Spotify AB 2021: 6) to more efficiently determine users' taste attributes by using technology (rather than relying on users to rate content). Although Spotify said in 2021 that they do not intend to use the system and that the idea is solely an asset in the form of intellectual property (Biometric Update 2021), this statement is telling. It points to a horizon of acceptability (hence why it was conceived and published) that it is also ethically problematic (hence the public relations communication), and it suggests that the idea is commercially valuable (even if Spotify does not pursue it, others will want to, and Spotify will still benefit from the sale of its asset).

Like Amazon's Alexa, this patent is solely audio based, involving analysis of content—namely, "speech recognition" and content metadata regarding "Emotional State, Gender, Age, [and] Accents" (Spotify AB 2021: 1). It also analyses environmental metadata, including "Physical Environment (i.e., Bus, Metro, Train, Outdoor, School, Coffee, Park...)" and "Social Environment (i.e., Alone, Small Group, Party...)" (Spotify AB 2021: 1). The patent notes that environment, for example, may comprise speech other than that of the Spotify user, including "sounds from vehicles on a street, other people talking, birds chirping, printers printing, and so on" (Spotify AB 2021: 7). Using environmental metadata to triangulate behaviour increases Spotify's confidence in the accuracy of its recommendations because Spotify would know the nature of the event, where it is, and with whom it is happening. The patent also exemplifies its profiling activities, detailing, "a listening and rating history and links to associated profiles such as those of the user's friends or colleagues, as well as storage for the user's existing music collection and/or library" (Spotify AB 2021: 9).

The profiling and targeting of users' emotions based on their voice biometrics and background noise, then, is a potential future for consumer-facing recommender systems. In Amazon's patent it is also envisaged as ubiquitous and in real-time, while in Spotify's patent it is also envisaged as encompassing who the user is with and where. With such services on the near horizon, and with pending European Union legislation suggesting that transparency about whether such systems are deployed allows people to make informed choices, it is imperative to understand whether people are willing to engage with such systems, and if so, on what terms. To that end, we stimulated discussion with diverse participants on potential benefits and concerns about these intimate means of profiling and targeting, as described below.

Workshop Participants

We conducted ten focus group workshops, each two hours long, across 2021. Our participants (n=46) were recruited through a professional research panel using previously submitted demographic information and an additional screener questionnaire. As age is the biggest indicator of differences in attitude towards emotional AI in the UK (as described earlier), we purposively sampled participants (Miles, Huberman, and Saldana 2014) to ensure that age-related differences were well represented, recruiting three younger (18–34 years old) groups (n=12) and three older (+65 years old) groups (n=13). This is a sampling strategy of *maximum variation* to expose the widest range of viewpoints.

Furthermore, to capture more diverse views we recruited participants for two groups composed of people who self-identify as disabled (n=10) and two groups belonging to UK ethnic minorities (n=11). Collecting data from these additional groups with protected characteristics is important because such groups have historically been ignored by AI developers (AI Now 2019; Benjamin 2019; Packin 2021). Emotional AI systems could have much to offer disabled people who may, for instance, be more reliant on technology to enable communication, socialising, and employment. However, they could also worsen ableist discrimination by, for instance, failing to categorise and distinguish the vast diversity of often invisible, temporally fluctuating health disabilities, and by prescribing a value-laden benchmark of what constitutes "normality" (AI Now 2019; Packin 2021). Indeed, more broadly, the rise of emotional AI has been accompanied by concerns about the accuracy of such systems, including concerns about coded bias in procedures for classifying emotions (algorithms); the simplistic universalising taxonomies of emotion expressions; and the inferences that can be drawn (Barrett et al. 2019; Stark and Hutson 2021). Furthermore, analyses of industrial applications of certain forms of biometric emotional AI have found these systems to be racist in failing to be calibrated for black skin or voices (Rhue 2018; Gal 2020; Koenecke et al. 2020). For instance, black speakers were twice as often misunderstood than white speakers in services by Amazon, Apple, Google, IBM, and Microsoft (Koenecke et al. 2020). Any concerns about emotional AI, then, would likely be *intensely manifested* among such groups, with disabled participants also sensitised to their benefits.

Conducting the focus group workshops online was necessary as COVID-19 social distancing restrictions were in place across 2021. The size of each of our ten workshops was small, typically four or five participants (to maximise their ability to speak) and lasted no more than two hours (to minimise online fatigue). Before data-collection, the research project was approved by the university's research ethics board,

and each participant was presented with an informed consent form and plain English information sheet about the project. These materials are publicly archived at UK Data Service along with the research protocol and full transcripts of the anonymised focus groups (Laffer 2022a).

Deploying an Interactive Narrative Method Using Design Fiction and ContraVision

Gauging attitudes towards emerging technologies among lay people is not straightforward due to the potential complexity of technologies, lack of clarity on how data are processed, inability to see potential privacy settings, and difficulty of situating abstract propositions (such as emotion profiling) in everyday life. To overcome these obstacles, we developed an interactive narrative method where we introduce and explore focus group topics through use of narrative forms involving interactive elements, such as branching narratives or in-text decisions, to give participants agency and opportunity to influence the story. We deployed the open-source storytelling tool, Twine, an interactive writing tool that enables simple construction and delivery of multimodal narratives (Laffer 2022b).

We drew from design fiction principles (Bleecker 2009). This is an approach that designs objects or technologies within a fictional world or narrative in order to explore not the technology itself, but the impact and influence of technology on people's lives, social institutions, and norms. Work in this field is varied, including not only physical design (Bleecker 2009) but also narrative forms (Jensen and Vistisen 2017; Markussen and Knutz 2013). As such, we created a fictional narrative world where emotional AI is situated in everyday circumstances. Our narrative documents a day-in-the-life of the main protagonist encountering mundane instances and affordances of different emotional AI and emotion profiling technologies. This is initiated by introducing "AffecTech," a fictional technology developer used to connect the disparate and unfamiliar examples of emotional AI that participants encountered within the narrative. Use cases were selected based on the likelihood of appearance in lived contexts, now and in the near future (for instance, informed by patents). As well as the voice-based emotional AI recommender systems discussed in this paper (the home-hub smart assistant and the music recommender system), the narrative also introduces emotional AI within a bus station surveillance sensor, a sales call evaluation and prompt tool, an emotoy, a hire-car automated system, and emotion profiling from social media algorithms.

In each workshop, all the participants (as a group) are talked through the interactive narrative by a moderator, where they are presented with these emotional AI use cases sequentially. Most use cases followed a sequence of (a) simple, neutral introduction of the technology and its emotion profiling; (b) a binary choice emerging from the technology use (e.g., to accept or reject the emotional AI technology's recommendations); and (c) a "ContraVision" element with a positive and negative event or outcome, after which they are asked to reflect on the use case and its harms and benefits. The narrative then moves onto the next use case.

The binary choices gave participants an opportunity to make a simple judgement call on their initial impressions of the technology, which prompted informed group discussion based on agreement and disagreement in views. It also provided a point of reflection and comparison when the ContraVision elements were introduced. "ContraVision" is the repeated presentation of a narrative showing the use of technology with competing positive and negative outcomes for the same scenario (Mancini et al. 2010). We deployed the ContraVision technique within the narrative across most of our use cases by presenting participants with a reasonably good outcome and a less good outcome, but we avoided dystopian and utopian hyperbole. As Mancini et al. (2010: 161) find, ContraVision elicits a wider spectrum of responses than a single presented perspective. They also suggest that ContraVision is particularly effective when "researchers have reason to believe that the technology is likely to raise subtle and elusive personal, cultural and social issues that can potentially jeopardize its adoption" (Mancini et al. 2010: 161).

In the fictional narrative, the Home-hub registers that the protagonist's voice is unhappy and sounds unwell. Reflecting the voice analytics functionality envisaged in the patent for Amazon's Alexa, the Home-hub voice assistant asks, "Would you like me to order something to make you feel better?," with participants responding "yes" or "no" (see Figure 1).

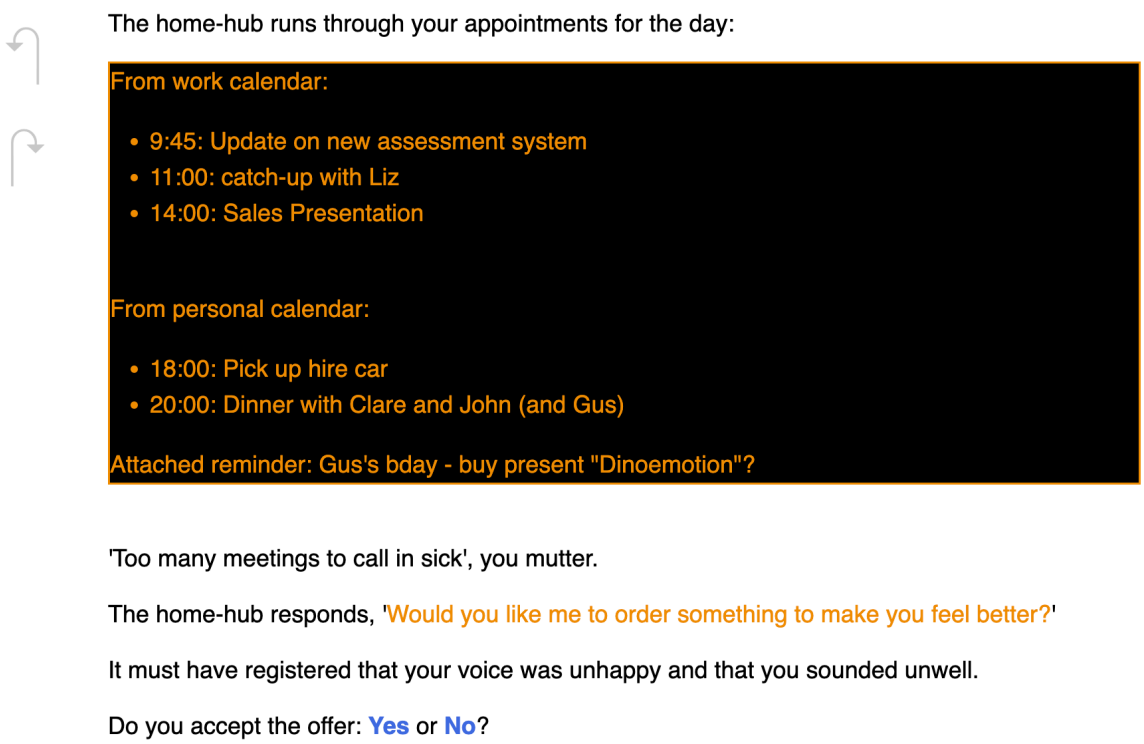

Figure 1: *Home-hub Use Case: Binary Choice*

Depending on their choice, participants would see one of two response variants (designed to preserve narrative sense), with the Home-hub suggesting the protagonist has cold symptoms that started two days ago, ordering (or offering to order) paracetamol and decongestant, and asking if they would like a reminder in their calendar to call a doctor if symptoms do not improve. Participants were then asked to reflect on these potential actions. In offering potential diagnostic and healthcare support (through purchasing paracetamol and scheduling help), this forms the positive ContraVision element of this use-case (see Figure 2).

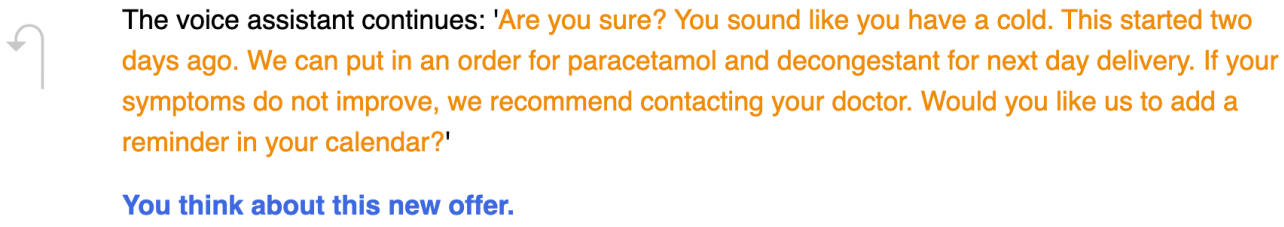

Figure 2: *Home-hub Use Case: Positive ContraVision*

The narrative progresses with the protagonist scrolling through their social media feed to notice two personalised ads. One is for a new brand of orange juice highlighting its health benefits. The other is for their local pharmacy. These ads form the negative part of the ContraVision aspect, as profit is being generated through profiled emotion data (see Figure 3).

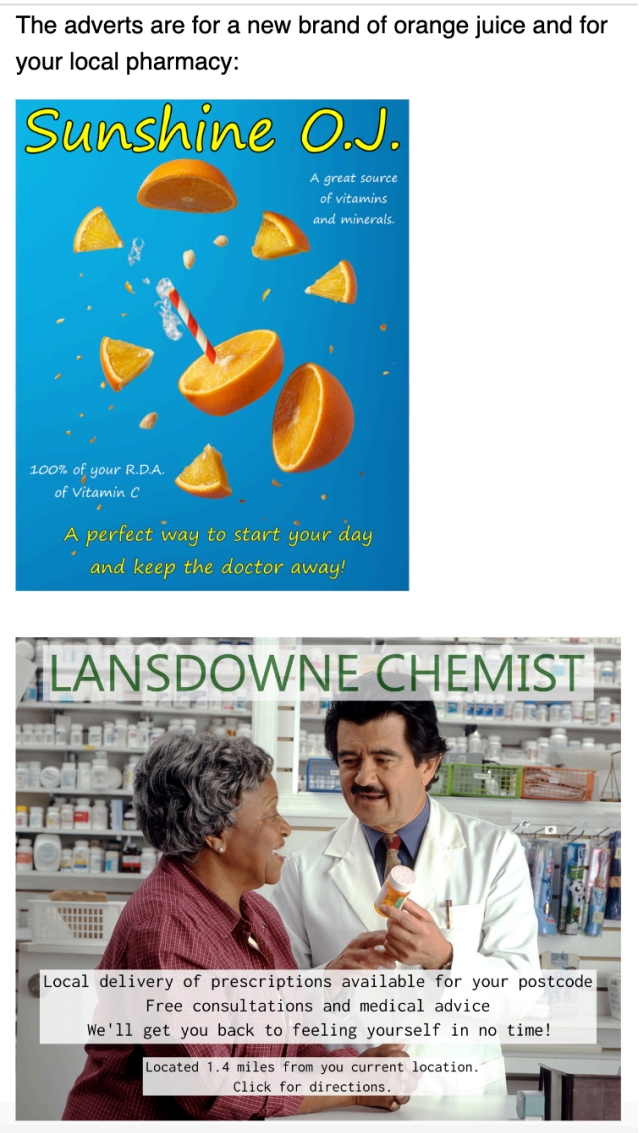

Figure 3: *Home-hub Use Case: Negative ContraVision*

After several other encounters with emotional AI, the protagonist boards a bus to work and decides to listen to music on their smartphone via Spotify. On opening Spotify, the protagonist is asked to review new Terms and Conditions, which we wrote to reflect key features of Spotify's patent: "These say that to improve their recommendations to you, they would like to collect data about emotional states through speech recognition, mood and social setting, by establishing where you are, and who you are with. Do you click: 'yes, happy to accept'/'no, thanks'?" (see Figure 4). Due to time constraints imposed by the workshop, and the need to progress the narrative onto other use cases without disturbing narrative flow, this particular use case does not feature the ContraVision elements of the other six use-cases explored in our workshop. Nonetheless, guided by the moderator, this short use case was effective in stimulating discussion of how the emotional AI functionality might be used (with positive and negative applications discussed), and in allowing participants to evaluate and express opinions.

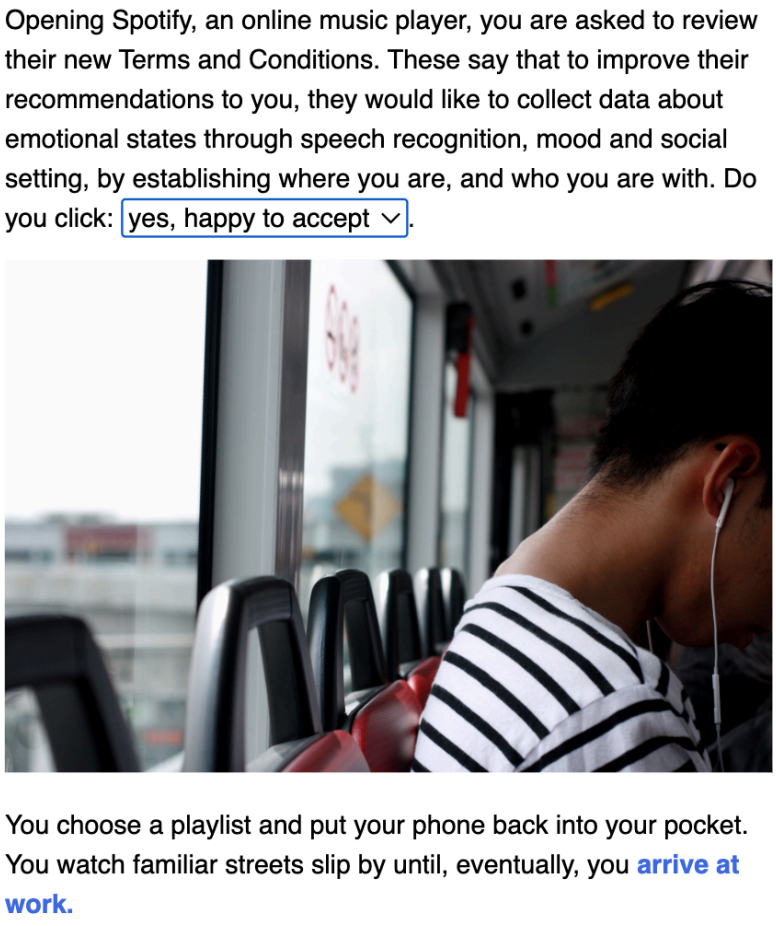

Figure 4: *Spotify Use Case: Binary Choice*

Throughout the workshop, we wanted to find the shortest route to helping participants understand how the emotional AI technologies might be deployed to enable an informed discussion about these proposed systems. To that end, this use case also refers to "Spotify" (rather than the fictional company "AffecTech") for quick recognition of functionality. While not naming Spotify might have been neater (by having all use case examples under the fictional company, "AffecTech"), the trade-off to naming Spotify was quicker comprehension by our participants (given Spotify's market dominance in music recommender systems). Conversely, we opted for "Home hub" rather than naming Amazon as there are several other well-known voice assistants (e.g., from Google and Apple) and we did not want to slow down, or divert, the workshops by prompting discussions about specific brands.

Throughout, we minimise social desirability bias (namely, the tendency of participants to answer questions in a manner that will be viewed favourably by others) by ensuring that the moderator kept the explanation of what was happening in each use case neutral and ensuring that ContraVision elements (both the written text and any accompanying visuals) presented both a good and a bad outcome of roughly equal valence, without celebratory or dystopic hyperbole.

The complete workshop narrative is available online,[1] and for a fuller methodological account of the development and implementation of this approach, see Laffer (2022b). In short, we found that, in addition to encouraging a diversity of responses, the advantages of our innovative approach included: increased engagement, avoiding participant fatigue and leading to rich data; mitigation of variation in participants' digital literacy; rapid familiarisation with abstract and unfamiliar topics; and a focus on people's lived experiences while exploring impacts of new technologies.

Coding Data

Data coding followed an adaptive approach (Layder 1998) balancing inductive insights from data with deductive theory. Utilising qualitative data analysis software, NVivo, a hand-coded approach was employed (Miles, Huberman, and Saldana 2014) to ensure sensitivity to the context of the workshops' discussions, initially annotating sentences and paragraphs of inductive interest. This surfaced data-first codes, some common to both use cases discussed in this paper (such as "control" and "functionality") and others that were use-case-specific. Reflecting on these codes, we abstracted them into broader themes informed by deductive interest in critiques of emotional AI, capitalism, and user-based studies on digital personal data. All authors undertook this process, reached similar conclusions, and then debated and agreed on four key themes (reported below). These key themes were found across all four of our participant types (namely, older [O], younger [Y], disabled people [D], and those of ethnic minority [E]) across each of our two use cases.

Key Findings

Our four key themes comprise "usefulness," "resignation," "uneasy terms of engagement," and "human-first." "Usefulness" covers a mix of positive and negative views, while the other three key themes are more negative.

Usefulness

Across all our participant categories, the Spotify use case (concerning Terms and Conditions to collect data about emotional states) was widely disliked and the Amazon Alexa styled Home-hub was not popular. However, where positive attitudes could be found they focussed on the functionality of being *useful* and relevant for social situations and to the self. For instance, in regards to Spotify, Paul (Y2) states, "there are parts of it that I could understand why it'd be great… like you pick 'read the room,' you could also do the same thing with mood." When the moderator suggests that one function could be recognising "that you're in a public place, so it turns the volume down or up, depending on how loud it is," disabled participant Penny responds, "That might be a good idea" (D2). With the Amazon Alexa styled Home-hub, our participants saw utility in an attentive AI, sensitive to context, who could discern emotional cues and disposition. For instance, "it might think I was slightly stressed or slightly under the weather and perhaps it could help" (Violet O3). Reflecting positively on its usefulness to others, Yasmine (E2) states, "I'm thinking for people who don't really have anyone that can emotionally attend to their needs."

Resignation, but Not by All

We use the term "resignation" to refer to acceptance of the technology, albeit stemming from a sense of powerlessness. In the Home-hub use case, we find significant evidence of resignation in existing attitudes towards profiling across all our participant groups. For example, Hannah (D2) states, "I don't think it would bother me too much because there's so many similar things, like when you go on Facebook or any kind of similar thing." Theo (E2) says, "We don't really know, but whether we like it or not, we're already impacted by artificial intelligence anyway." Alice (Y3) articulates a desire for control in an environment where the net effect is to create a sense of resignation through constant behaviour profiling. She makes the simple but telling observation that, "It'd be nice to have the option to turn it off. Like, okay, you know, if people want a bit more privacy... I don't think you can do it, but then to say, 'oh, I'm going to turn this off for a few

[1] https://eaitwine.neocities.org/

days.'" This belies very low expectations of environments created for participants, as well as the view that technology is something that happens *to* them, not *with* or *for* them.

In the Spotify use case, on being asked to directly consider their choice about whether to accept new Terms and Conditions for emotion profiling, resignation is evident in that older and younger participants alike admit that: "I never read them. I just go, 'Yeah, go on'" (Linda O3). Luis (Y1) elaborates: "most of those playlists, we select them based off of how we're feeling at the moment, or what we would like to hear. So in one sense, they're already collecting that emotional data, without they realizing it."

Interestingly, resignation was not apparent among ethnic and disabled participants in this use case. Rather, they stressed the necessity of being able to choose how to interact with the platform on data collection and processing. For instance, on being asked if she would be happier with the scenario if the data were processed locally in her phone or speaker rather than centrally by Spotify, Emily (D2) responds: "If you can choose which you want, that's fine to me because I'm probably at some point going to want both but not always." Rejecting Spotify's Terms and Conditions, Maira (E2) states: "I'd click, 'No, thanks.' I don't like the idea of them collecting data about my emotional states, mood and social setting. I just think it is just far too much."

Uneasy Terms of Engagement

Related to the theme of "resignation," but conceptually distinct, this theme about the "uneasy terms of engagement" with data extractive industries surfaces where participants see the goal of profiling as inevitable, trust-breaching and pervasive, or diverging from existing functionality. This took form of concerns about the *data sought and stored* and *function creep*.

Participants from all four of our categories expressed concern about the *data sought and stored* by Spotify and the Home-hub, and the wider digital environment oriented around extracting data. In the Home-hub use case, Joanne (O2) flags the role of training data in machine learning: "For AI to improve, it's got to learn… by collecting the data to build up its understanding; so, I can't believe that it's going to collect it and then get rid of it." Phillip (O1) articulates a perceived breach of trust between person and affect-sensitive cloud-based synthetic personalities: "feels like a personal thing between you and Alexa, but actually she's picking up on information, and that information is going off to Amazon, or Google, or whatever, and you're not necessarily aware of that." There is also a more general disquiet about pervasive surveillance. With the Home-hub use case, Carol (Y3) reveals: "it does definitely make me feel uncomfortable and it's a reason why I don't have Alexa because of, it listens to all of your data and the surrounding things." In the Spotify use case, Patrick (D2) states that recommending mood-sensitive music "could be a good thing, but like I said, it could also be a bad thing because it's another way of monitoring you."

On *function creep*, the Spotify use case is criticised across all our participant categories. Recollecting that in the fictional narrative, Spotify seeks to access data about emotional states through speech recognition, mood, and social setting—by establishing where you are and who you are with—Harry (O3) states, "I just want to listen to music. What I eat, these other things they want me to tell them about, have nothing to do with them." Tom (Y2) asks: "why do they need to be able to see and hear me if I'm just listening to music?" In context of emotion-enhanced services, Theo (E2) asks whether Spotify is still just: "an online music player? Now, you're changing your functionality, you're changing what you're there for. You're collecting data, emotional states. You're no longer what you were all about. All you were about is about music."

Human-First

Across our diverse groups we found a wish for a "human-first" environment. This is one that attaches importance to people, including a belief in people's *agency*, their right to *not be replaced* by a machine, and the *inability of AI systems to correctly interpret human emotions and how people live*. We see this as especially important in that the finding reflects European Union data and AI policy agendas, but it also speaks to a more fundamental sense of alienation. This is simultaneously humanities-based and political.

The first is about tension between a strong sense of subjectivity in relation to algorithmic recommender media environments and algorithmic prescription; the second is the more Marxist observation that digital emotional labour is reified into commodity relations where data about emotion come to stand against the individual.

For both use cases, a recurring statement across all our participant groups is that they do not want loss of *human agency*. With Spotify, they do not want an algorithm to select music for them, especially when based on emotion profiling. Nigella (O2), for example, says, "I wouldn't use it, because […] I don't need some artificial intelligence advising what I need to what I listen to. I know that from within myself what I feel like listening to." We found her "from within myself" remark telling in that it potentially belies disquiet with a world already shaped by algorithmic decision-making. This tips into alienation in that one feels separation from what one has little choice to contribute to. Indeed, with the Home-hub use case, across all backgrounds sampled, there was a view that the emerging socio-technical world was not of their choosing. This is exemplified well by Tim's (O1) comment on the biometric expansion of existing algorithmic services: "All I see is, in front of me, a machine that's come from, I don't know which company, it could be Amazon, could be a rival, whatever, but it's actually now acquired massively enhanced opportunities to push this to the frontier [and] that's what's frightening me."

Others, especially older participants, express concern about *human replacement*, observing that the Home-hub is akin to being nagged by your parent or partner. Amanda (O2) states, "It's a bit like having your mum there telling you… not allowing you to think for yourself." In relation to the claimed affective and empathic qualities of these systems, Tim (O1) observes that "the machine pretends to be your friend because it's got all these empathetic characteristics, but really it is not your friend at all."

Across all our participant categories, our "human-first" theme also takes shape in discussion about the nature of decision-making and the *inability of AI systems to correctly interpret human emotions and how people live*. For instance, on the Home-hub use case, Patricia (Y1) remarks, "I think humans are very complex and I don't know how well an app or a device could read their emotions and order something that will genuinely make them happy." Tarun (E2) observes: "there might be some misinterpretation, that you're in a bit of a rush, and so it just misinterpreted what you're doing."

Discussion

Distilling Industrial Ambition on Emotional AI

Given the centrality of emotion to decision-making (Barfar 2019), extraction of emotion data is of keen interest to surveillance capitalism (McStay 2018; Zuboff 2015, 2019). We have exemplified this in patents by Amazon and Spotify, which detail their desire to ingest biometric data. This includes, in the case of Amazon, real-time data about the emotional and physical characteristics of users based on biometrics (voice analytics), as well as more standard behavioural and contextual targeting criteria. In the case of Spotify, it includes biometric (voice-based) data about users' emotional state, as well as their gender, age, accents, physical and social environment, listening and rating history, and links to profiles of friends or colleagues. As noted earlier, Spotify has since announced that it does not intend to use the system, but it can still benefit from sale of its patented asset. Given ambition (or at least, testing of the waters and accrual of commercial value in patents) by platforms to use biometrics to further enable their profiled targeting, the turn towards emotional AI is in plain sight. The key feature of such industrial ambition in consumer-facing recommender systems is the expansion of the bandwidth of data collection to encompass biometrics, enabling delivery of ever more emotionally targeted, personalised content under the humanising guise of empathic systems, where, in the words of one of our participants, "the machine pretends to be your friend… but really it is not your friend at all" (Tim O1).

Situating Users' Views on Emotional AI

Our qualitative study was designed to surface informed views from diverse users, to explore if people are willing to engage with emotional AI in various contexts, and if so, on what terms. When asked to consider the prospect of emotion profiling in consumer-oriented platforms deploying recommender systems such as the Amazon Alexa-like Home-hub and Spotify, participants from all four of our categories could appreciate its "usefulness." This included its relevance for social situations (in being able to emotionally read a room), for others (to attend to people's otherwise unmet emotional needs), and for themselves (in being able to alter the environment to compensate for disability). Emotional AI in consumer-facing recommender systems, then, has appeal for some, if genuinely useful. This helps answer the "why" question arising from previous British-based surveys on users' views towards biometric-based emotional AI technologies that find a small majority preference for use in consumer-facing products and services (Bakir and McStay 2022; McStay and Rosner 2021).

Previous user-based studies find that people have issues controlling their personal data online, desiring more control (Hartman et al. 2020) but not expecting to get it (Dencik and Cable 2017; Draper and Turow 2019; Hargittai and Marwick 2016). This is evident in our data in the theme of "resignation," namely acceptance of the technology, but accompanied by feelings of powerlessness. This appears in both our use cases to different extents. It appears across all four participant categories in the case of Home-hub, but it only appears among younger and older participants in the case of Spotify where, instead of voicing resignation, our ethnic and disabled participants stress the necessity of being able to *choose* how to interact with platforms on data collection and processing. There has been increasing popular attention in recent years to explicit and implicit biases (especially racial, but also ableist) in algorithms on platforms, with well-publicised examples of algorithmic racial discrimination on online home-sharing company Airbnb, ride-sharing services Uber and Lyft, social media platform Facebook's advertising, and search engine Google's results (Benjamin 2019; Noble 2018; Turner Lee 2018; Velkova and Kaun 2021). As such, our ethnic minority and disabled participants may be sensitised to such issues. They may even have experienced such discrimination given racial and ableist bias in the technologies, or they may feel already too surveilled given racial bias in society (AI Now 2019; Benjamin 2019) (although this is not articulated). Their attitude aligns more with studies that find that users value the possibility of shaping recommendation algorithms (Harambam et al 2019; Shin 2020), and in some cases, actively try to do so to develop, resist, or disrupt algorithmic outputs (Cotter 2018; Karakayali, Kostem, and Galip 2018; Velkova and Kaun 2021).

The related theme of "uneasy terms of engagement" with data extractive industries surfaces as participants from all four of our categories express concern about the data sought and stored in the Spotify and Home-hub use cases, seeing the goal of such profiling as inevitable, trust-breaching, and pervasive. The Spotify use case is criticised across all our participant categories on surveillant grounds of function creep. While previous studies find an "algorithm knowledge gap" even in digitally literate societies (Zarouali, Nelberger, and de Vreese 2021), our findings suggest that many of our participants are aware of the drawbacks of datafied surveillance, at least as far as consumer-facing recommender systems are concerned. Also, the surveillant theme of function creep overlaps with data protection concerns of data minimisation and purpose limitation principles, as Art. 5(1)(c) of the European Union 's General Data Protection Directive requires that personal data be adequate, relevant, and limited to what is necessary in relation to the purposes for which it is processed. Indeed, in addition to technical questions of what data are genuinely required to fulfil personalised service offers (Biega et al. 2020), our findings on strategic changes in function suggest the need for closer attention to how service purposes are defined and whether amendments to purposes are acceptable when users have invested effort in curating the service (e.g., creation of playlists). This is particularly important given that there is only "marginal regulatory and judicial guidance on the interpretation of data minimization" (Biega et al. 2020).

Across our diverse groups we also find a wish for a "human-first*"* environment. This theme is one that attaches importance to people (including their *agency*, with older people in particular expressing that *people should not be replaced by a machine*). Although as researchers we were sensitive to post-phenomenological insistence on attention to the specifics of technology (Verbeek 2011), the need to avoid encouraging

celebratory or dystopian responses, and the need to situate these technologies in everyday settings, our participants nonetheless express a keen sense of generalised *alienation*. This stands against theorisation by surveillance scholars that, rather than being alienated, users would likely be cultivated into being docile surveillant subjects by the usefulness of consumer surveillance (as, for instance, West [2019] argues regarding Amazon's Alexa). Rather, our user-based study finds alienation, with dimensions not only concerning choice and agency but also an estrangement from emotion-based digital products that participants would co-produce, a feeling of a lack of control, and a sense that the world is not their own but rather someone else's system. An additional facet to alienation is that of being at odds with algorithmic stupidity. Indeed, in the Home-hub use case, all our participant categories noted the *inability of AI systems to correctly interpret human emotions and how people live*. This accords with longstanding philosophical critique of the ability of statistics and algorithms to make decisions that humans might make, as well as more recent studies querying baselines, thresholds, and ground truths of emotion in emotional AI systems (McStay 2018, 2023), with many more studies criticising the accuracy and in-built biases of such systems (Noble 2018; Turner Lee 2018; Barrett et al. 2019; Koenecke et al. 2020; Stark and Hutson 2021). Also, with Home-hub, across all backgrounds sampled, there was a view that the emerging socio-technical world was not of their choosing, indicating a desire for more control over their environment, which should be oriented foremost towards humans rather than machines. This finding accords with emerging literature on people's views towards AI. For instance, a US study finds student support for an AI teaching assistant *as long as* it is useful and easy to communicate with (Kim et al. 2020); and a Canadian study finds that elderly people can develop trust in autonomous homecare systems *if* they believe they can operate it and find a sense of belonging and feelings of social interactivity from the AI system (Shareef et al. 2021).

Conclusion

Our previous quantitative work into British users' views on emotional AI in various near horizon use cases, including consumer-facing areas, finds disquiet from various groups. Younger people are more open to such technologies if care is taken over personal identification, but most older people dislike this premise (Bakir and McStay 2022). British parents have ambivalent views on the prospect of toys that utilise children's emotions, but they express disquiet over its intrusiveness and what happens to the data (McStay and Rosner 2021). Shedding light on reasons underlying such disquiet, our novel qualitative approach with diverse users finds some support for such technologies for their "usefulness," but we primarily evidence contemporary "human-first" concerns. These insights provide a richer understanding of whether emotional AI technologies should be deployed at all in consumer-facing recommender systems, and if so, on what terms (summarised below).

Policy and Developer Recommendations

Our study enables policymakers and designers of recommender and emotional AI systems to better understand the lifeworlds of users and incorporate insights in future products, policies, and regulations. Platforms using biometric-based emotional AI in recommender systems, in particular, should pay attention to the desire by participants across all our diverse categories for the products to be genuinely useful and relevant for social situations, for others and for themselves. Policymakers should be alert to our participants' desire for a human-first environment that flags a generalised sense of alienation. They should also be alert to the strong desire expressed by those who identify as ethnic minorities or disabled to be able to *choose* how to interact with platforms in term of the collection and processing of data about their emotions (which some would likely refuse); and they should be aware of the concerns expressed by all four participant categories about the uneasy terms of engagement, especially concerning data sought and stored and about function creep. As the European Union's draft regulation on AI continues to be shaped, we question whether having heightened transparency obligations on automated emotion recognition systems and biometric categorisation disclosure is enough to allow people to make informed choices or step back from a given situation (European Commission 2021: 14). More transparency without an accompanying sense of being able to make a real choice on the matter will only fuel resignation, already common among our participants. Having qualitatively considered diverse users' views on emotional AI in consumer-facing recommender

systems, we conclude that they signify a sense that technological and algorithmic environments are things that happen *to* people rather than *for* them. This is not simply a matter of trust or lack of understanding, but a long-standing awareness that uses of new technologies do not always make life better.

Limitations and Future Research

Our qualitative study is limited in that it is not generalisable to wider populations, especially as our participant categories were purposively sampled to unearth diverse views concerning emotional AI technologies. Furthermore, while we have taken efforts to minimise social desirability bias, this cannot be ruled out. Finally, our user-based findings (especially the more reflective, critical findings) may be a product of our research design, where participants are asked to actively consider the otherwise transparent or hidden practices and purposes emerging from digital devices and services: in real-life situations, such critical concerns may not arise.

Nonetheless, our qualitative workshops, in which diverse British-based participants reflect upon near-horizon use cases of emerging technologies in a fictional narrative world, allow participants to better understand how these technologies may be used and bring their informed views to the foreground in empirical detail, which future surveys can test across broader populations. In addition, qualitative studies with non-British-based publics (for instance, in parts of the world with lower digital literacies and with looser regulations on, and cultures of, data protection and AI, such as much of the Global South) would also be useful, as most of the studies reported in this paper (with the exception of Siles et al.'s [2020] study of Costa Rica) are based on populations from the Global North (British, Dutch, American, Canadian, and Korean). Finally, as emotional AI systems start to be rolled out across the world, we recommend studies on the actual lived experiences of their users.

Acknowledgments

This project was funded by the UK's Economic & Social Research Council (Grant ref. ES/T00696X/1).

References

AI Now. 2019. *Disability, Bias, and AI*. https://ainowinstitute.org/disabilitybiasai-2019.pdf.

Andrejevic, Mark. 2011. The Work that Affective Economics Does. *Cultural Studies* 25 (4–5): 604–620.

Bakir, Vian, and Andrew McStay. 2018. Fake News and the Economy of Emotions: Problems, Causes, Solutions. *Digital Journalism* 6 (2): 154–175.

———. 2022. *Optimising Emotions, Incubating Falsehoods: How to Protect the Global Civic Body from Disinformation and Misinformation.* Cham, CH: Palgrave Macmillan.

Ball, Kirstie. 2009. Exposure: Exploring the Subject of Surveillance. *Information, Communication and Society* 12 (5): 639–657.

———. 2019. Review of Zuboff's The Age of Surveillance Capitalism. *Surveillance & Society* 17 (1/2): 252–256.

Barfar, Arash. 2019. Cognitive and Affective Responses to Political Disinformation in Facebook. *Computers in Human Behavior* 101: 173–179.

Barrett, Lisa Feldman, Ralph Adolphs, Stacey 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* 20 (1): 1–68.

Benjamin, Ruha. 2019. *Race after Technology: Abolitionist Tools for the New Jim Code.* Cambridge, UK: Polity.

Berardi, Franco. 2009. *The Soul at Work: From Alienation to Autonomy*. Los Angeles, CA: Semiotext(e).

Biega, Asia J., Peter Potash, Hal Daumé III, Fernando Diaz, and Michèle Finck. 2020. Operationalizing the Legal Principle of Data Minimization for Personalization. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), China, July 25–30*. New York: Association for Computing Machinery.

Biometric Update. 2021. Privacy Advocates, Spotify Only See the Bad in Each Other. May 7. https://www.biometricupdate.com/202105/privacy-advocates-spotify-only-see-the-bad-in-each-other [accessed May 3, 2023].

Bleecker, Julian. 2009. Design Fiction: A Short Essay on Design, Science, Fact and Fiction. Near Future Laboratory (blog), March 17. https://blog.nearfuturelaboratory.com/2009/03/17/design-fiction-a-short-essay-on-design-science-fact-and-fiction/ [accessed May 3, 2023].

Cotter, Kelley. 2018. Playing the Visibility Game: How Digital Influencers and Algorithms Negotiate Influence on Instagram. *New Media & Society* 21 (4): 895–913.

Dencik, Lina, and Jonathan Cable. 2017. The Advent of Surveillance Realism: Public Opinion and Activist Responses to the Snowden Leaks. *International Journal of Communication* 11 (19): 763–781.

Draper, Nora A., and Joseph Turow. 2019. The Corporate Cultivation of Digital Resignation. *New Media & Society* 21 (8): 1824–1839.

Eslami, Motahhare, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I "Like" It, Then I Hide It: Folk Theories of Social Feeds. In *Proceedings of the 34rd Annual SIGCHI Conf.Human Factors in Computing Systems 2016, San Jose, CA, May 7–12*, 2371–2382. New York: Association for Computing Machinery.

European Commission. 2021. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. Brussels, COM (2021) 206 final 2021/0106 (COD), April 21. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence.

Gal, Danit. 2020. China's Approach to AI Ethics. In *The AI Powered State: China's Approach to Public Sector Innovation* edited by Hessy Elliot, 53–62. London: Nesta. https://media.nesta.org.uk/documents/Nesta_TheAIPoweredState_2020.pdf.

Gregory, Karen, and Jathan Sadowski. 2021. Biopolitical Platforms: The Perverse Virtues of Digital Labour. *Journal of Cultural Economy* 14 (6): 662–674.

Harambam, Jaron, Dimitrios Bountouridis, Mykola Makhortykh, and Joris van Hoboken. 2019. Designing for the Better by Taking Users into Account: A Qualitative Evaluation of User Control Mechanisms in (News) Recommender Systems. In *Thirteenth ACM Conference on Recommender Systems*, *Copenhagen, Denmark, September 16–20,* 69–77. New York: Association for Computing Machinery.

Hardt, Michael, and Antonio Negri. 2000. *Empire*. Cambridge, MA: Harvard University Press.

Hargittai, Eszter, and Alice Marwick. 2016. "What Can I Really Do?": Explaining the Privacy Paradox with Online Apathy. *International Journal of Communication* 10: 3737–3757.

Hartman, Todd, Helen Kennedy, Robin Steedman, and Rhianne Jones. 2020. Public Perceptions of Good Data Management: Findings from a UK-Based Survey. *Big Data & Society* 7 (1): 1–16.

Iliadis, Andrew, and Amelia Acker. 2022. The Seer and the Seen: Surveying Palantir's Surveillance Platform. *The Information Society* 38 (5). https://doi.org/10.1080/01972243.2022.2100851.

Jensen, Thessa, and Peter Vistisen. 2017. Ethical Design Fiction: Between Storytelling and World Building. In *ETHICOMP 2017 Conference Proceedings: Values in Emerging Science and Technology, Turin, Italy, June 5–8,* 1–14. Aalborg, DK: Aalborg University.

Jin, Huafeng, and Shuo Wang. 2018. Voice-based Determination of Physical and Emotional Characteristics of Users. US Patent US10096319B1, filed March 13, 2014, and issued October 12, 2018. Google Patents. https://patents.google.com/patent/US10096319B1/en.

Karakayali, Nedim, Burc Kostem, and Idil Galip. 2018. Recommendation Systems as Technologies of the Self: Algorithmic Control and the Formation of Music Taste. *Theory, Culture & Society* 35 (2): 3–24.

Kienscherf, Markus. 2022. Surveillance Capital and Post-Fordist Accumulation: Towards a Critical Political Economy of Surveillance-for-Profit. *Surveillance & Society* 20 (1): 18–29.

Kim, Jihyun, Kelly Merrill, Kun Xu, and Deanna D. Sellnow. 2020. My Teacher is a Machine: Understanding Students' Perceptions of AI Teaching Assistants in Online Education. *International Journal of Human–Computer Interaction* 36 (20): 1902–1911.

Koenecke, Allison, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion, Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky, and Sharad Goel. 2020. Racial Disparities in Automated Speech Recognition. *Proceedings of the National Academy of Sciences* 117 (14): 7684–7689.

Laffer, Alexander. 2022a. Attitudes Towards Emotional Artificial Intelligence Use: Transcripts of Citizen Workshops Collected Using an Innovative Narrative Approach, 2021. Data collection, UK Data Service, SN: 855688. https://dx.doi.org/10.5255/UKDA-SN-855688.

———. 2022b. Using an Online Narrative Approach to Explore Diverse Participants' Understanding of Emerging Technology: Citizen's Perspectives on Living With Emotional AI. *SAGE Research Methods: Doing Research Online*, March 1. https://doi.org/10.4135/9781529604122.

Layder, Derek. 1998. *Sociological Practice: Linking Theory and Social Research*. London: Sage.

Lazzarato, Maurizio. 2014. *Signs and Machines: Capitalism and the Promotion of Subjectivity*. Los Angeles, CA: Semiotext(e).

Mancini, Clara, Yvonne Rogers, Arosha K. Bandara, Tony Coe, Lukasz Jedrzejczyk, Adam N. Joinson, Blaine A. Price, Keerthi Thomas, and Bashar Nuseibeh. 2010. ContraVision: Exploring Users' Reactions to Futuristic Technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, Georgia, April 10–15*, 153–162. New York: Association for Computing Machinary.

Markussen, Thomas, and Eva Knutz. 2013. The Poetics of Design Fiction. In the *DPPI '13: Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, Newcastle upon Tyne, UK, September 3–5*, 231–240. New York: Association for Computing Machinary.

McStay, Andrew. 2018. *Emotional AI: The Rise of Empathic Media*. London: Sage.

———. 2023. *Automating Empathy: Decoding Technologies That Gauge Intimate Life*. New York: Oxford University Press.

McStay, Andrew, and Gilad Rosner. 2021. Emotional Artificial Intelligence in Children's Toys and Devices: Ethics, Governance and Practical Remedies. *Big Data & Society* 8 (1). https://doi.org/10.1177/2053951721994877.

Miles, Matthew B., A. Michael Huberman, and Johnny Saldana. 2014. *Qualitative Data Analysis*. London: Sage.

Neville, Stephen J. 2020. Eavesmining: A Critical Audit of the Amazon Echo and Alexa Conditions of Use. *Surveillance & Society* 18 (3): 343–356.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Packin, Nizan Geslevich. 2021. Disability Discrimination Using AI Systems, Social Media and Digital Platforms: Can We Disable Digital Bias?" *SSRN*, January 11. https://ssrn.com/abstract=3724556.

Rhue, Lauren. 2018. Racial Influence on Automated Perceptions of Emotions. SSRN, December 17. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281765.

Sadowski, Jathan. 2019. When Data Is Capital: Datafication, Accumulation, and Extraction. *Big Data & Society* 6 (1): 1–12.

Shareef, Mahmud Akhter, Vinod Kumar, Yogesh K. Dwivedi, Uma Kumar, Muhammad Shakaib Akram, and Raman Ramakrishnan. 2021. A New Health Care System Enabled by Machine Intelligence: Elderly People's Trust or Losing Self Control. *Technological Forecasting and Social Change* 162. https://doi.org/10.1016/j.techfore.2020.120334.

Shin, Donghee. 2020. How Do Users Interact with Algorithm Recommender Systems? The Interaction of Users, Algorithms, and Performance. *Computers in Human Behavior* 109. https://doi.org/10.1016/j.chb.2020.106344.

Siles, Ignacio, Andrés Segura-Castillo, Ricardo Solís, and Monica Sancho. 2020. Folk Theories of Algorithmic Recommendations on Spotify: Enacting Data Assemblages in the Global South. *Big Data & Society* 7 (1). https://doi.org/10.1177/2053951720923377.

Spotify AB. 2021. Identification of Taste Attributes from an Audio Signal. US Patent 10,891,948, filed February 21, 2018, and issued January 12, 2021. https://patentimages.storage.googleapis.com/5c/0f/84/8b53c2903a82ba/US10891948.pdf.

Stark, Luke, and Jevan Hutson. 2021. Physiognomic Artificial Intelligence. SSRN, September 24. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3927300.

Stiegler, Bernard. 2010. *For a New Critique of Political Economy*. Cambridge, UK: Polity.

Turner Lee, Nicol. 2018. Detecting Racial Bias in Algorithms and Machine Learning. *Journal of Information, Communication and Ethics in Society* 16 (3): 252–260.

Velkova, Julia, and Anne Kaun. 2021. Algorithmic Resistance: Media Practices and the Politics of Repair. *Information, Communication & Society* 24 (4): 523–540.

Verbeek, Peter-Paul. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago, IL: University of Chicago Press.

Von Otterlo, Martijn. 2014. Automated Experimentation in Walden 3.0: The Next Step in Profiling, Predicting, Control and Surveillance. *Surveillance & Society* 12 (2): 255–272.

West, Emily. 2019. Amazon: Surveillance as a Service. *Surveillance & Society* 17 (1/2): 27–33.

Zarouali, Brahim, Natali Helberger, and Claes H. de Vreese. 2021. Investigating Algorithmic Misconceptions in a Media Context: Source of a New Digital Divide? *Media and Communication* 9 (4): 134–144.

Zuboff, Shoshana. 2015. Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology* 30 (1): 75–89.

———. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Pow*er. New York: PublicAffairs Books.

Zuboff, Shoshana, Norma Möllers, David Murakami Wood, and David Lyon. 2019. Surveillance Capitalism: An Interview with Shoshana Zuboff. *Surveillance & Society* 17 (1/2): 257–266.