



Article

The Infopolitics of feeling: How race and disability are configured in Emotion Recognition Technology

new media & society

1–19

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14614448241235914

journals.sagepub.com/home/nms



Kerry McInerney 

University of Cambridge, UK

Os Keyes

University of Washington, USA

Abstract

In this article, we argue that facial emotion recognition technology (facial ERT) reproduces historical forms of pseudoscience based on the concept of quantifiable and unequally distributed emotional capacity. Drawing on Kyla Schuller's *Biopolitics of Feeling* and Colin Koopman's theory of infopower, we put forward the term 'the infopolitics of feeling' to describe how facial ERT encodes culturally 'correct' or normative forms of emotional expression that have historically been used to define and delineate what it means to be human. To make this argument, we provide a close reading of *Girl Decoded*, the autobiography of Rana el Kaliouby, the founder and former CEO of the leading Emotion artificial intelligence (AI) firm Affectiva. *Girl Decoded*, we argue pits el Kaliouby herself – portrayed as the empathetic, liberal, emotionally expressive and ideal 'feeling' subject – against two non-normative figures: the unfeeling autistic and the inscrutable Oriental who must be 'cured' through Affectiva's facial ERT.

Keywords

Affect, Affectiva, AI ethics, artificial intelligence, autism, biopower, emotion, Emotion AI, emotion recognition technology, facial recognition, oriental inscrutability, Orientalism, pseudoscience, race science, racism, sentimentalism

Corresponding author:

Kerry McInerney, The Leverhulme Centre for the Future of Intelligence, University of Cambridge, 16 Mill Lane, Cambridge CB2 1RX, UK.

Email: kam83@cam.ac.uk

Introduction

In her 2020 biography *Girl Decoded*, Rana el Kaliouby, the founder and CEO of the then-leading Emotion artificial intelligence (AI) firm Affectiva writes,

Artificial Intelligence, or AI, is the science of training computers to think and reason like human beings. Emotion AI is focused on training computers to recognize, quantify, and respond to human emotion, something that traditional computers were not built to do. My goal is not to build emotive computers, but to enable human beings to retain our humanity when we are in the cyber world. This book – my life – is about the quest to humanize technology before it dehumanizes us (el Kaliouby and Colman, 2020).

Emotion, in el Kaliouby's worldview, is positioned as a central, if not defining, characteristic of what it means to be human. Consequently, she justifies her foray into the field of Emotion AI by claiming that technology – understood as cold, emotionless and unfeeling – must be 'humanized' through emotion, lest it dehumanise us. Emotion AI refers to a wide range of technologies that aim to record, measure, interpret and categorise human emotional states by drawing on a heterogeneous set of data sources (including facial expressions, voice, heart rate and other signals) as well as attempts to simulate human emotion using AI. el Kaliouby's excitement about developments in Emotion AI and the field of affective computing is reflected in a rapidly growing Emotion AI market. For example, in 2022, *Analytics Insight* predicted that the emotional AI market would rise to US\$37.1 billion by 2026 (Akash, 2022). This market growth has occurred in spite of intensifying critiques of Emotion AI firms' fundamental premise: that human emotions can be readily identified from the face across geographical and cultural contexts, and effectively interpreted by machines (Barrett et al., 2019). Consequently, critics have argued that Emotion AI may replicate historical forms of pseudoscience and scientific racism, such as phrenology and physiognomy, and even give them new legitimacy in the form of 'objective' AI tools (AI Now, 2018; Atanasoski and Vora, 2019; y Arcas et al., 2023). Indeed, Stark and Hutson (2022: 932) include Emotion AI in their definition of 'physiognomic AI . . . using computer software and related systems to infer or create hierarchies of an individual's . . . perceived character, capabilities, and future social outcomes based on their physical or behavioral characteristics'.

However, while existing critiques of Emotion AI systems rightly question the pseudoscientific premises of this field of inquiry, they rarely engage with how Emotion AI firms and founders like Rana El Kaliouby frame emotion (and subsequently, their products) in relation to what it means to be human. While Cave (2020), Elam (2022) and y Arcas et al. (2023), among others, have all importantly explored how forms of racial pseudoscience, such as IQ testing, physiognomy and phrenology shape the field of AI, there has been insufficient scrutiny of how the false sciences associated with understanding, measuring and labelling emotional states affect contemporary AI development and deployment (for an exception, see Atanasoski and Vora, 2019). In this article, we focus on a specific subsection of Emotion AI: facial emotion recognition technology (facial ERT), which aims to read, measure and deduce expressions from people's faces. While a pseudoscientific logic may underlie the broader field of Emotion AI, we argue that facial ERT in particular reanimates historical forms of scientific racism that use the concept of emotion – and

specifically, culturally ‘correct’ or normative forms of emotional expression – to define and delineate what it means to be human. In particular, these tools reproduce a narrow definition of liberal personhood premised on the idea that one’s capacity to feel emotions and sense pain according to a predefined normative standard are central to what it means to be human. Drawing on the histories of affect and race science put forward by scholars, such as Kyla Schuller and Xine Yao, we suggest that facial ERT represents the newest frontier of such race science, where the political project of trying to grant machines emotional intelligence further codifies sexist, racist and ableist hierarchies of the ‘human’.

We coin the term ‘infopolitics of feeling’ to describe how facial ERT represents the latest vector of emotional governance and racial pseudoscience. The infopolitics of feeling combines Kyla Schuller’s (2018) critical analysis of the 19th-century sciences of emotion in the *Biopolitics of Feeling* with Colin Koopman’s concept of ‘infopolitics’. Koopman uses the term ‘infopolitics’ to gesture towards a much longer genealogy of information as a ‘technology of power’ (Koopman, 2018: 105). We use the term ‘infopolitics’ to connote how data that are collected about someone’s emotion state or capacity for emotional expression – no matter how spurious – is used for two purposes: (1) to actively govern and control people whose emotional expression or capacity for emotion is deemed non-normative, and compel them to feel in ways that are ‘correct’ or ‘right’ and (2) to (re)produce a figuration of normative liberal personhood where being human is contingent on being able to feel in the ‘right’ ways. Rather than merely seeing facial ERT tools as fallible products that are unable to achieve their aims, we instead see them as part of a broader infopolitical project that compounds and entrenches existing forms of discrimination and dehumanisation.

In the first section of this article, we examine the historical role (or lack of) that emotion has played in computing, and then lay the groundwork for understanding the contemporary field of facial ERT. We then examine the infopolitical connotations of facial ERT technologies by exploring the history of emotion and how the concept of emotional capacity has historically been used to create taxonomies of humanity and justify dominative colonial, racist and ableist political projects. Next, we turn to the central case study of this article: the autobiography of Rana el Kaliouby, the founder of the affect recognition firm Affectiva, to examine how disability, race and gender are configured throughout her work. We discuss how el Kaliouby’s autobiography epitomises the ‘infopolitics of feeling’ described earlier in the article through a critical analysis of how she represents two groups that have historically been racialised and discriminated against due to their supposed ‘incapacity’ to feel or to express emotion in correct and socially acceptable ways: autistic people and East Asian people, whose faces and character have been broadly constructed in the Western racial imagination as ‘inscrutable’ to the Western eye.

Facial ERT and the politics of feeling ‘right’

How do you feel about emotion recognition?

Emotion has historically been portrayed as outside the scope of computing. Notwithstanding the long countercurrent of research demonstrating the role emotion plays in the conceptualisation and design of computing (Keyes and Austin, 2022; Su

et al., 2021), theorists and practitioners have treated emotion as not only uninteresting to developers, but also as a problem that should be eradicated. This absence of feeling – and the implicit distinction between rationality and emotion that gives rise to it – has been a central part of both the promise and power of computing. By this, we mean both that the development and deployment of computing technology has often been motivated by the promise that a machinic substitute will produce less fallible, biased and contextual outcomes. The cultural cachet of computing is also deeply entangled with (masculine) epistemic ideals of emotionless rationality and notions of legitimacy (see Adam, 2006; Code, 1991).

As computing has entered realms more explicitly coded as part of ‘the social’, however, an alternative view has arisen, commonly labelled *affective computing*. This term gained prominence with the work of Rosalind Picard, whose 2000 book of the same name argues that emotion is not only not contrary to rationality, but a central part of it (Picard, 2000: 247). Correspondingly, the optimal computer is not a device lacking emotion, but rather a device that can, well, *compute* it: recognise it, analyse it and integrate it into its response to the user. Partly driven by Picard herself, the affective computing movement – built around the concept of ERT – is often portrayed as both a more human and more humane form of artificial intelligence. While technology is frequently interpreted as antithetical to feeling, emotion and care, the field of facial ERT aims to create tools that can accurately detect human emotions and label human affective states (Boyer, 2015). Founders of companies like Affectiva claim that facial ERT helps facilitate human-AI experiences by training AI to recognise and generate human-like emotions, thus allowing for more streamlined integration into human decision-making processes in everyday life.

This shift towards facial ERT has not been without controversy – and for good reason. In alignment with broader trends in computational critique, particularly AI ethics, researchers have highlighted the epistemic assumptions and limitations of facial ERT in both theory and practice. Many researchers emphasise the lack of verifiable scientific evidence that someone’s emotional state can be successfully identified from their face (Stark and Hoey, 2021). In a systematic review of the existing psychological literature on the topic, Feldman Barrett et al. (2019) found that how people communicate Ekman’s six basic emotions vary substantially across different situations, cultures and people. Their work implies that emotion recognition technologies may struggle to correctly identify emotions due to their lack of contextual knowledge and that training an algorithm to be able to understand a particular emotion may be a difficult, if not completely impossible, task. Facial ERT also raises significant questions around surveillance and the extent of personal biometric data collection, as evidenced by regulatory moves against workplace data collection (Kak and West, 2023; McStay, 2020; McStay and Rosner, 2021; Mantello and Ho, 2023; Podoletz, 2023).

Even more importantly, scholars like Drage and Mackereth (2022), and organisations like AI Now (2018) emphasise that facial ERT is grounded in very little empirical scientific evidence and thus may reanimate historical forms of pseudoscience, such as physiognomy and phrenology. After all, the core premise underlying ERT based on facial data is that internal emotional states map squarely onto external appearance. On one hand, this ignores the ordinary practice of deception; people often school their faces in such a

way that does not reflect how they really feel at any given moment in time. On the other hand, it also reveals a more insidious logic, suggesting that the ‘truth’ of someone’s thoughts, feelings and emotions can be accurately read from their external appearance. As Wendy Chun (2009: 10) argues,

race in these circumstances was wielded – and is still wielded – as an invaluable mapping tool, a means by which origins and boundaries are simultaneously traced and constructed and through which the visible traces of the body are tied to allegedly innate invisible characteristics.

From here, it is only a short jump to the phrenological and physiognomic assumption that external appearance is a reliable indicator of the character. We can see this assumption operating in several malicious uses of facial recognition technology for scientifically racist purposes, such as algorithms designed to deduce criminality from the face or identify someone’s sexual orientation (y Arcas et al., 2023). While these pseudoscientific algorithms certainly cannot perform the tasks they claim to do, their mere existence replicates the foundational work conducted by foundational figures in the field of physiognomy and phrenology, such as Cesare Lombroso, the Italian scientist who believed that criminality was inherited and that this criminality was reflected in criminals’ head shape (y Arcas et al., 2023). As Michele Elam notes, the drive to categorise, taxonomise and classify people and faces in the fields of AI and machine learning evokes (and is perhaps the latest iteration in) the Enlightenment compulsion to create a colonial order of things (Elam, 2022).

From Biopolitics, to Infopower, to the Infopolitics of feeling

Taken together, the critiques of facial ERT we explored above highlight valuable concerns with the idea of affective computing in general as a panacea to problems of computing’s coldness. What we wish to draw attention to and focus on is instead the biopolitical and *infopolitical* implications of EAI – whether it is perfected or not – and the use of emotion as a ‘practice of governance’; as part of ‘the forms of reason and organisation through which individuals and groups coordinate their various activities, and the practices of freedom by which they act within these systems, following the rules of the game or striving to modify them’ (Tully, 2002: 538). Examinations of biopolitics – defined here as ‘the state-centred exercise of the power to “foster” or “disallow” human life . . . through regulation of the biological “life” of a population’ (Diprose and Ziarek, 2018: 8) – have often focussed on rationality in general, and scientific reason in particular, as its main driving force. This includes examinations of the shift between biopolitics and what Colin Koopman (2018) refers to as ‘infopower’; the increasing use of information as a (putatively self-contained) source of power and control. Much as with biopower, many analyses of infopower focus on the rationalist aspects of information; on the deployment of putatively neutral and logical informatic systems to classify, distinguish, clump and control (Cheney- Lippold, 2011; Elam, 2022). This is not to say that critical scholarship ignores emotion; some, such as data feminists Catherine D’Ignazio and Lauren Klein (2020), argue for emotional and embodied knowledge to be incorporated into technology design. But as this suggests, the focus for such scholarship often

remains – much as it does for Picard and colleagues – on the idea of emotion as a solution to informatic injustices, and to infopower.

The problem is that in both biopolitics and infopolitics, cold rationality is not the only ‘form of reason’ deployed; to the contrary, emotion has often been used *itself* as a tool of power. A growing body of scholarship examines how care, usually posited as a feminist ethic or principle, has simultaneously functioned as a mode of biopolitical governmentality (Anderson, 2021; Gagen, 2015; Murphy, 2019; Semel, 2022; Stevenson, 2014; Ticktin, 2011; Puig de la Bellacasa, 2017; van Dooren 2014). Kyla Schuller, in her thoughtful and incisive *The Biopolitics of Feeling*, has highlighted the ways in which conceptions of emotion (and their connections to forms of life and personhood) regularly appeared as a tool of governance and regulation in the 19th and early 20th centuries. She investigates how ‘nineteenth-century biopower consolidated in a sentimental mode that regulated the circulation of feeling throughout the population and delineated differential relational capacities of matter, and therefore the potential for evolutionary progress, as the modern concepts of race, sex, and species’ (Schuller, 2018: 2). In 19th-century US culture the perceived *impressibility* of the human body – the ability of the body to affect and be affected by its external environment – was used to distinguish ‘civilised’ bodies from ‘primitive’ ones (Schuller, 2018: 5). While the White body was characterised as impressible, sensitive and progressive, racialised bodies were cast as insensate, impulsive and incapable of evolutionary change (Schuller, 2018: 4). Crucially, these notions of impressibility and emotional capacity were used to stratify people into different types of life, distinguishing between those with the capacity to care and those who could not feel properly, and thus could only be ‘felt for’ rather than with.

And this phenomenon is not a historical one. These gendered and racialised perceptions of bodily sensitivity and impressibility have had violent and profound consequences; this includes the brutal exploitation of indigenous labour in Australia, whose bodies were considered less sensitive by White Australian settlers, through to the ongoing racist assumption that Black people do not feel pain as intensely as White people do in US medicine (Phillips, 2015). Schuller (2018: 2) traces the history of sentimental biopower to contextualise the contemporary biopolitical role of emotion, and specifically how ‘white feelings, in the context of the United States, are the fertile products of racialized vulnerability, disposability, and death’. Meanwhile, Sara Ahmed (2014) states that emotions ‘work to shape the ‘surfaces’ of individual and collective bodies’, in ways not easily reducible to the claim that feelings are good (or bad) and highly tied up in existing ideas of what feelings, expressed by whom, and in response to what, are acceptable (Ahmed, 2014: 1). The moral valence of feelings, in other words, is heavily tied into (and reinforces) biopolitical frameworks of power. This can also be seen in ongoing social movements, particularly those – such as feminist social groups – that seek in part to recuperate the validity of emotion as a form of expression and knowledge. The ethnographic work of Sarita Srivastava (2005) demonstrates how this recuperation enables the deployment of emotion by (White) feminists within these spaces as a way of warding off and invalidating concerns about racism from activists of colour. Emotion may be conceived in contrast to ‘cold rationality’ and the power that has accrued to it. But as historic and present examples demonstrate, emotion is hardly an escape from power itself.

As a result, much contemporary feminist and critical race scholarship complicates the simplistic coding of emotion as a feminist form of knowledge, instead highlighting how feeling, affects, and emotion all function as vectors through which gendered and racialised relations of power operate. In Schuller's (2018: 2) words, it shows how 'sentimentalism, in the midst of its feminized ethic of emotional identification, operates as a fundamental mechanism of biopower'. Emotions play a critical role in constructing and categorising gendered and racialised bodies, as the affects that create, stick to, and shape the body and its surfaces help transform 'what is "lower" or "higher" into bodily traits' (Ahmed, 2014: 4). These power relations shape who is allowed to express emotion; which forms of emotional expression are considered to be legitimate; and whose forms of emotional expression are rendered unrecognisable as emotion. As Xine Yao (2021) notes, anti-racist protest slogans, such as 'white tears, white fragility, white women's tears, white men's tears' all foreground how structural Whiteness operates through the affective fragility of White people when confronted with experiences of race and racism. Meanwhile, work by scholars, such as Leslie Bow (2022) demonstrates how White people's relationships with racialised people are often characterised by the queasy relations of fetishisation and fear; racialised people are frequently desired, pitied, hated, commodified, loved and despised by White people simultaneously, and positioned as the passive object of this range of emotions rather than as the rightful holder of them.

Affectiveness, impressibility and emotion more broadly thus play a central role in the production of gendered, racialised and ableist hierarchies (Atanasoski and Vora, 2019; Stark and Hoey, 2021). Yet, despite important critiques of certain applications of ERT for both its pseudoscientific premises and its replication and of sexist and racist relations of power, there has been little sustained investigation into how facial ERT intersects with the use of feeling as a qualification for being human.¹ If emotion is used as a category for defining who counts and who does not count as human, relegating some people to the status of infrahuman or inhuman, we must seriously grapple with how facial ERT builds on biopolitical histories of feeling to create a new *infopolitics* of feeling. This should not surprise us, given that Koopman (p. 168) defines infopower not as a successor to biopower, but as something 'deposited on, or layered on, the sediment of earlier strata of power'. Given the logics of power at work in feeling, we seek to explore how this appears in computing efforts to confront and integrate emotion.

Methodology

In this article, we examine the infopolitics of facial ERT through a close reading of the autobiography of Affectiva founder Rana el Kaliouby. Based on el Kaliouby's doctoral and postdoctoral research at the MIT Media Lab's Affective Computing group, Affectiva's software has been a leading product in the field of affective recognition technology since the company's inception; indeed, in her autobiography el Kaliouby credits Affectiva's team for popularising the term Emotion AI. el Kaliouby has been recognised as a leading woman in technology and, in 2019, as one of the BBC's Hundred Women (BBC News, 2019). In 2021, the Swedish firm Smart Eye acquired Affectiva for US\$73.5 million, and el Kaliouby became Smart Eye's deputy CEO (O'Brien, 2021). In 2021, el Kaliouby

released her memoir, *Girl Decoded: A Scientist's Quest to Reclaim our Humanity by Bringing Emotional Intelligence to Technology*. The autobiography follows el Kaliouby through her childhood as a 'nice Egyptian girl', to her experiences as a computer science PhD student in Cambridge; the origins of her research in facial ERT as an attempt to help autistic children learn how to understand and express emotions; her postdoctoral research on ERT at the MIT Media lab; the founding of her facial ERT startup company Affectiva; and the widespread application of Affectiva software and various use cases across the world.

el Kaliouby's memoir provides vital insight into the beliefs and motivations of one of the leading entrepreneurs and computer scientists in the field of facial ERT. It is thus ripe for a critical analysis regarding how el Kaliouby perceives emotion and its relationship to the figure of the human, as well as how her beliefs reflect and permeate the field of facial ERT more broadly. Affectiva is not, of course, representative of the entire field of facial ERT, nor is el Kaliouby its only or definitive spokesperson. In our critical analysis, we do not want to imply that *Girl Decoded* represents a 'ground truth' of either el Kaliouby or Affectiva's core principles. Corporate memoirs like *Girl Decoded* function as performances that aim to promote a particular political agenda relating to the societal function of tech firms, and, relatedly, the leading role-played by tech innovators and CEOs like el Kaliouby herself. In this sense, *Girl Decoded* follows in the footsteps of tech entrepreneurs like Steve Jobs, Mark Zuckerberg and Elon Musk by personifying Affectiva through el Kaliouby and her personal narrative of growth, entrepreneurship and (self-)discovery. However, given Affectiva's leading role in the field of facial ERT, we interpret *Girl Decoded* – as both entrepreneurial bildungsroman and corporate messaging – as a useful index for the wider infopolitics of feeling that underpins the field of facial ERT. *Girl Decoded* encapsulates the role that el Kaliouby thinks that Affectiva – and, we infer, facial ERT as a technology – can and should play in contemporary US society.

This article in no way aims to invalidate el Kaliouby's undeniable achievements, especially given the gendered barriers and expectations levied against her as she built her career as a leading computer scientist and entrepreneur. Rather, we point towards the limits of representational politics by demonstrating how el Kaliouby's approach to emotion and facial ERT is fundamentally shaped by gendered, racialised and ableist hierarchies that humanise some at the expense of others. Through our analysis of her text, we ask: how does the el Kaliouby presented in *Girl Decoded* theorise and conceive of emotion, and how do they structure or reinforce ideas about whose feelings (and in what form) are legitimate and which are not? We focus on two particular figures deemed emblematic of gendered and racialised unfeeling in this article, partly due to their sociocultural pervasiveness, and partly due to their re-animation in technological form through the spectre of facial ERT. The first is the ableist portrayal of autism as the inability to properly feel, understand, process or express emotion, which is then used to establish control over autistic people and their bodies in the name of ableist saviourism. The second is the trope of the 'inscrutable Oriental', which frames 'Oriental' people broadly construed as opaque and emotionally unreadable to the Western eye.

Girl decoded and the infopolitics of feeling

Girl decoded and the liberal, feeling subject

Throughout *Girl Decoded*, el Kaliouby emphasises time and time again how emotion is absolutely fundamental to what makes us human. In the introduction, titled ‘Emotion Blind’, she opens her autobiography with a story of a group of teenagers who watched a man called Jamel Dunn drowning and, instead of helping him, they recorded the entire incident on their phones. el Kaliouby and Colman (2020) frame Dunn’s death as a symptom of a societal ‘empathy crisis’, arguing that

everyday, we encounter people who display a similarly shocking lack of empathy, not to mention basic civility . . . we, as a society, are in increasingly dangerous territory: we are at risk of undermining the very traits that make us human in the first place.

el Kaliouby and Colman (2020) argue that while intolerance and cruelty are not new to the social media age, they are strongly amplified by the advent of ‘emotion blind’ technologies which, according to el Kaliouby, dehumanise those behind the screen and make it ‘easy to forget that we are talking to and about other human beings’

In her discussion of the ‘empathy deficit’, el Kaliouby and Colman (2020) refer to ‘genocide, mass killings and slavery’ as ‘stains on our past’ that ‘still plague us today’. This universalising framing of genocides, mass murder and enslavement flattens the complex political, social, economic and historical causes behind these different phenomena into one single, driving factor: ‘empathy deficit’. It is also a problem to which el Kaliouby conveniently offers a technological solution: facial ERT, which she claims is ‘part of the cure’ (el Kaliouby and Colman, 2020). el Kaliouby and Colman’s (2020) choice of the word ‘cure’ here is no accident: she goes on to state that facial ERT can help heal or ‘repair the damage’ caused by conducting our lives within a digital, ‘emotion-free zone’. This is the cause to which el Kaliouby dedicates her early work, to the extent that she frames herself as fighting on behalf of and for ‘humanity’. For example, when presenting her PhD research to her research group for the first time, el Kaliouby and Colman (2020) write that their colleagues saw ‘the lack of emotion, the ‘clear-eyed’ calculated objectivity of a computer’ as what made them more effective than humans. Consequently, el Kaliouby and Colman (2020) argue that she had to win them over by building ‘a strong case for humanity’.

el Kaliouby’s framing of emotion as the final frontier of what it means to be human, and the need to protect this humanity against the cold, emotionless spectre of cyberspace, reproduces the liberal humanist framing of ‘feeling’ as a distinctly human capacity, and one that is integral to full liberal personhood. In doing so, she builds on the forms of algorithmic emotional management and ‘empathy hacking’ that is becoming increasingly influential in the tech industry. Take, for example, the use of virtual reality (VR) to generate empathy and develop a greater range of emotional reasoning, a project that Lisa Nakamura refers to as ‘virtuous VR’. Nakamura: 48, 54 argues that virtuous VR companies, which position VR as an ‘empathy machine that connects people across difference’, offer a new form of identity tourism for the 21st century. This latest iteration of online identity tourism stretches beyond the temporary and recreational assumption of ‘exotic’

gendered and racialised avatars, extending to the emotional occupation of a foreign body or a humanitarian victim (Nakamura 2020). And ‘virtuous VR’ is not the only example of how the tech industry attempts to synthetically generate empathy and ‘hack’ or coerce the body into feeling ‘right’. Cynthia Bennett and Daniela Rosner (2019: 2) highlight how empathy-building activities in the field of human-computer interaction ‘diminish disabled perspectives, separate the roles of disabled people and designers, and stage the disabled experience as a spectacle’ by recentering able-bodied designers and their own emotional growth at the heart of the design process.

Here, the use of technology to govern and regulate emotional expression, and ensure that participants express emotion in the service of greater ‘humanity’ and ‘compassion’ bears a striking resemblance to el Kaliouby’s insistence that facial ERT can help prevent the forms of violence and genocide that occur in the absence of feeling. Like Bennett and Rosner’s critique of empathy-building activities for HCI designers and Nakamura’s critique of virtuous VR, our approach to facial ERT as an empathetic and empathy-building tool does not malign the reality of human suffering or the importance of compassion. Instead, we take aim at how el Kaliouby’s framing of facial ERT as the response to the ‘shocking lack of empathy’ that shapes contemporary societies entrenches historical power disparities between the emotional, empathetic liberal humanist subject and its object of pity. In addition, despite her grand claims to create technologies that will serve all of humanity, Affectiva’s products are primarily focussed on integrating facial ERT into cars and creating effective advertising services, alongside other equally profitable commercial applications (el Kaliouby and Colman, 2020), highlighting the ways that ERT, as Jeff Nagy (2022) highlights, has served to make disability ‘a rhetorical, conceptual, and material resource for the expansion of surveillance capitalism’.

The failure of *Girl Decoded* to meaningfully challenge the power relations that underpin tech-generated forms of empathy lies, in part, with the book’s individualist ideology. *Girl Decoded* narrates the story of Affectiva and the development of facial ERT as a kind of corporate bildungsroman, where the development of the software is intimately twinned with el Kaliouby’s own social and emotional growth in a distinctly US narrative of liberal self-realisation. In the book’s opening, el Kaliouby and Colman (2020) write,

in striving to become the ‘expert’ I needed to be in human emotion in order to teach machines about emotion, I found myself turning the spotlight on my own emotional life . . . Ultimately, decoding myself – learning to express my own emotions and act on them – was the biggest challenge of all . . . my work and my personal story are inseparable; each flows into the other. And so this book is a chronicle of that dual journey – the quest to equip machines with EQ and, in the process, unlock my own EQ.

Thus, *Girl Decoded* explicitly frames Affectiva’s growth as a story of el Kaliouby’s personal transformation into a fully fledged, feeling, liberal subject.

Crucially, el Kaliouby and Colman (2020) also frame this emotional growth narrative as part of a larger liberal success story, that of American multiculturalism. Central to el Kaliouby and Colman’s (2020) character arc is her transformation from a ‘nice Egyptian girl’ to a major technology CEO and a US citizen ‘thriving on the energy, vitality, and entrepreneurial spirit of this great country’.

Men and women of all nationalities, religions and backgrounds now bound together as American citizens. My eyes filled with tears; it was the official beginning of my new life as an American, an Egyptian American who has become part of this amazing mix of cultures united by a common ideal of freedom, opportunity, and democracy. Here is the place where you can bring your crazy idea and attempt to change the world, a place where risk-taking is admired, and where pushing boundaries is encouraged, and is deeply ingrained in the American consciousness.

While el Kaliouby does emphasise a multicultural Egyptian-American hybrid identity, *Girl Decoded* falls into well-established liberal tropes about the utopian success of the American ‘melting pot’, a narrative that serves to mask and erase the deep racial and gendered divides that continue to structure US society. We in no way want to undermine the personal and systemic sexism and Islamophobia el Kaliouby undoubtedly encountered, as both a computer scientist and a tech CEO. Nonetheless, *Girl Decoded*’s blend of memoir, bildungsroman, and corporate performance piece links her educational journey with her political self-actualisation as the empathetic subject idealised by the liberal West (Bastani, 2020). Moreover, el Kaliouby’s individualistic narrative aligns with what Catherine Rottenberg (2018: 5) has identified as the rise of neoliberal feminism, a variant of feminism with a distinctly ‘individualizing and political anesthetizing effect’. As Rottenberg (2018: 7) writes, neoliberalism’s ongoing and relentless conversion of all aspects of our world into ‘specks’ of capital, including human beings themselves, produces subjects who are individualized, entrepreneurial, and self-investing; they are also cast as entirely responsible for their own self-care and well-being. el Kaliouby represents the near-perfect neoliberal feminist subject: she is deeply entrepreneurial, individual, self-investing, and committed to the American dream of a perfectly multicultural yet entirely atomised and self-sufficient society.

Girl decoded and autism

If el Kaliouby represents the ideal subject, dead-centre on the scale of empathy, who represents the extremes? *Girl Decoded* makes clear that for one of those extremes, the answer is ‘autistic people’, her view of whom is central to not only el Kaliouby’s work, but the existence and flourishing of Affectiva as a company and Emotion Recognition as a domain.

Her actual understanding and depiction of autism mirrors that of Simon Baron Cohen, an (in)famous Cambridge neuroscientist who specialises in autism research. To Baron-Cohen (2002), autism is the manifestation of an ‘extreme male brain’. Humanity exists on a spectrum from most to least emotionally and collectively sensitive, a spectrum that maps to gender. Women are sensitive, empathetic; men logical, contained. Autistic people – characterised by him as having a ‘striking poverty’ of empathy and communicability – are thus masculinity taken too far: so independent and self-contained they (we) are locked in. Baron-Cohen has been consistently critiqued, with researchers not only highlighting the limited (and contradictory) empirical evidence for his views, but the way those views presume a strongly gendered normative ideal of human behaviour (Gernsbacher and Yergeau, 2019; Lockhart, 2020; Yergeau, 2013). Despite this,

Baron-Cohen's worldview (and its deficit-oriented premise) remains popular. That el Kaliouby mirrors these views (indeed, quotes them extensively) is unsurprising: she credits Baron-Cohen as her primary source of insight into autism, one who 'altered how I viewed the world and my work' (el Kaliouby and Colman, 2020: 101–102), having worked with him directly at Cambridge, and she states that, only a few days prior to meeting him, she had 'never heard of autism' (el Kaliouby and Colman, 2020: 91–92). To el Kaliouby and Colman (2020: 101–102), too, autism is a gendered insufficiency, one that is fundamentally masculine, and so that, predominantly impacts men and boys.

The insufficiency and lack of personhood of autistic people in el Kaliouby's mind could be the topic of an article in and of itself; at every point, el Kaliouby and Colman (2020: 91–92) manage to hit cultural tropes about autism, from portraying autistic people as exclusively male (and exclusively children, and childlike), to learning about autistic subjectivity solely through the perspectives of non-autistic relatives and self-described experts. Each is tired (and tiresome), but what is interesting is the image of (el Kaliouby's view of) autistic people that comes through as a result of their combination: an image we might call 'correctable insufficiency'. Autistic people are too male; too closed-off; too invulnerable to the feelings of others. Luckily, there is a fix – a *cure* – in the form of facial ERT, which serves to 'habilitate' (see Kim, 2017) the autistic person to normative sociality. A cure made more plausible given the autistic person's status, to el Kaliouby, as a child: as someone flexible, mutable and therefore moveable. More plausible as a result of her interpretation of Baron-Cohen's theories, which she summarises as teaching her that 'where an individual lands on the spectrum is not static' and that facial ERT could be used as an 'emotional prosthetic' (el Kaliouby and Colman, 2020: 102) by those struggling with emotional responsiveness and recognition.

el Kaliouby's analysis of autistic people, and the 'solution' to our existences, carries uncanny echoes of the gendered and racial aspects of biopolitical regimes of feeling. As Schuller discusses, the 19th and 20th century subject was frequently constructed and classified in relation to their sensitivity and responsiveness to the feelings of others (particularly in the case of the 'civilised' subject). For both conservative and liberal scholars of race and gender in that era, subjects were at risk of both under- and over-responsiveness – the former treated as masculine and the latter as feminine (Schuller, 2018: 16). The solution was 'sentimentalism', which 'worked to position the body's differential capacity of feeling as the object and method of state power . . . through the stimulation and regulation of the body's vital capacities' (Schuller, 2018: 20)

Like those 19th- and 20th-century scientists, el Kaliouby (and Baron-Cohen before her) portrays perceived autistic asociality, or lack of feeling, as masculine. And like those scientists, the solution is stimulation; the 'cure' of autistic existence through facial ERT, which works to regularise and normalise autistic relations to the world. In taking this approach, el Kaliouby positions autists as one extreme of a familiar biopolitical spectrum of feeling: the masculine, 'cut off' from sociality and relation, and in need of cure to approach a normal (Western) medium. el Kaliouby and Colman's (2020) framing of facial ERT as the 'cure' for an empathy deficit is particularly troubling given that she consistently describes this societal tech-based empathy deficit as a form of electronic autism: 'when it comes to the digital world, our computers have trained us to behave as if we lived in a world dominated by autism, where none of us can read one another's

emotional cues' This harmful stereotype runs throughout the whole of el Kaliouby's autobiography, positioning autistic people outside of the figure of the human (defined as those who can feel and express emotion 'correctly', according to societal norms). Those suffering from an empathy deficit – which, el Kaliouby implicitly suggests, includes 'monstrous' figures, such as perpetrators of genocide, mass killings and enslavers, as well as autists – are cast as inherently inhuman unless they are able to be 'taught' or 'trained' to feel correctly.

Girl decoded and the 'inscrutable Oriental'

A spectrum, of course, has two ends not one, and if autism and autistic people represent those capable only of deep, internal 'unfeelingness', who is the opposite? Whose feelings are too opaque and inscrutable to count as a liberal feeling subject? Who feels, but cannot express that emotion 'properly', and thus sits outside the category of the human? We now turn to consider a second trope of 'unfeelingness', albeit one that pivots from internally held feelings to externally connoted expression: the stereotype, applied to people of East Asian descent, of the 'inscrutable Oriental'. The exoticising and deeply racially laden term 'Oriental', once used to describe people and places located in the 'East' in opposition to the 'Occident', or 'West', is no longer acceptable in the US context.² Nonetheless, its racial history and associated structures of feeling remain in place, to the extent that Yao (2021: 171) writes that Oriental inscrutability is 'perhaps the most coherent racialized mode of unfeeling, the fact that it has a particular name indicating a structurally pervasive and lingering phenomenon in the Western cultural imagination'. Yao traces how this stereotype was enshrined in US culture and law by the US immigration apparatus and how the perceived inaffectability of Chinese people played a central role in the justification and the passing of the Chinese Exclusion Act (1882). Chinese people were excluded from the liberal, sentimental model of the human due to their supposed inability (or perhaps unwillingness) to emote in line with Western cultural norms; 'these unassimilable people', Yao (2021: 176) writes, 'are what the theorist Sara Ahmed would call affect aliens, to the extent that they are literalized as extraterrestrial'. The racialisation of Chinese people as 'affective aliens' is intimately related to the stereotypical association of Asian Americans with robots, machines, and computational intelligence (Bui, 2022; He, 2022; Huang, 2019; Roh et al., 2015; Shah, 2019 (Bow, 2022; Cheng, 2019; Sohn, 2008). Just as computers are considered antithetical to human forms of emotional intelligence, Chinese people and Asian Americans are constituted as affectively opaque and 'machine-like' in their supposed lack of emotional expression.

Girl Decoded similarly casts Chinese people as machine-like to a degree in their supposed inability to emote correctly. However, unlike autistic people, who she believes are unable to read and understand the emotions of others, el Kaliouby bemoans how Chinese people's emotions *cannot be read* by Affectiva's software. In the chapter 'Going Global', el Kaliouby details how Affectiva's software originally did not work for the Chinese market, threatening Affectiva's growth and el Kaliouby's personal quest to make people more emotionally transparent. While 'Going Global' is only a small chapter of Affectiva's

story, el Kaliouby's framing of Chinese people's emotional expression as non-normative builds on a much longer history of the racialisation of Chinese people as inscrutable.

Throughout 'Going Global', el Kaliouby interprets the software's inability to read Chinese faces as either a technical failure on the part of the algorithm or a problem with Chinese social norms regarding emotional expression. el Kaliouby describes how she 'fixed' the problem of Chinese emoting by ensuring that Chinese people used the software on their own, rather than in the presence of other people, in order to remove the social expectations surrounding 'correct' emotional expression. She also compelled the company to upload and include far more photographs of her impression of the more common baseline expression of Chinese people – what she calls the 'politeness smile' – to help the software distinguish between a smile of politeness and true happiness. The solution, according to *Girl Decoded*, is to 'fix' the algorithm through the addition of more data and also to 'fix' the emotional expression of Chinese participants by removing other participants from the room. (el Kaliouby and Colman (2020) argue that in countries, such as China or India, 'where group goals supersede those of the individual', people are more likely to hide or mask their emotions, 'especially negative emotions such as anger and contempt'. These emotions, she writes, are considered 'self-indulgent' (el Kaliouby and Colman, 2020). el Kaliouby and Colman (2020) write that

many of the Chinese test subjects wore a smile as their baseline expression; an ever-so-slight lip corner pull . . . It was the smile of politeness that I had often used myself as the 'nice Egyptian girl', the smile of a man or woman who didn't want to offend anyone, the play-it-safe smile'.

The Chinese subject is, in other words, not only emotionally repressed by non-Western cultural norms from which el Kaliouby has liberated herself; she is also troublingly opaque to the eye of the machine.

To truly work, *Girl Decoded* suggests, Affectiva's software must learn how to break the reserved, passive facade of the Chinese face. While Affectiva aims to teach autistic people how to properly read emotions, it aims to strip back the layers of deception and opacity that cloud Chinese people's ability to emote according to the software's particular needs. In doing so, *Girl Decoded* plays into the characterisation of Chinese people – and specifically, the Chinese face – as deceptive or mask-like. Arthur Smith's *Chinese Characteristics* (1890), the most read US text on China in the early 1900s, takes the 'usual expressionless visage' of the Chinese face as the defining characteristic of Chinese people, while the US travel writer Bayard Taylor decided that Chinese people had 'dull faces, without expression', producing an 'unconquerable aversion' on his part (Yao, 2021: 180). Likewise, Danielle Wong (2017: 40) writes that 'the Asian/American face has been, and continues to be, read as an inauthentic surface associated with both the machine and the mask'. The perceived disaffectedness of the Chinese face is deeply rooted in histories of racial pseudoscience, for in 19th-century physiognomic practice

the faculty of secretiveness is signaled by the degree that one's nostrils resembles those of a Chinese, for they are 'the most remarkable people in the world for secretiveness' – a point illustrated with an engraving of a generic East Asian face (Yao, 2021: 180).

Technologies like Affectiva threaten to continue the association of the Chinese face – and, physiognomically, the Chinese character – with emotional deception.

According to *Girl Decoded*, by adding more ‘polite smiles’ to Affectiva’s database, the algorithm learns to distinguish the politeness smile from one of ‘genuine happiness’. Affectiva’s perceived unmasking of the passive and docile Chinese face reflects the gendered tenets of Orientalism, where the ‘East’ is portrayed as feminine, alluring, and passive in comparison to a vigorous, active, and masculine ‘West’. While el Kaliouby draws on Baron-Cohen’s gendered framing of autism as a product of the hyper-systemising ‘extreme male brain’, Chinese people are positioned as both hyper-feminine yet also queerly non-normative in their emotional expression (Huang, 2022). Unlike the ideal liberal female subject in *Girl Decoded*, who demonstrates the right kind of feminine feeling, Chinese women function as a symbol of the guardedness and closed off nature of Chinese communities (and, in the terms of 19th-century retrograde race science, the Chinese ‘race’) (Yao, 2021). These gendered dynamics are further reinforced by *Girl Decoded*’s descriptions of unmasking the feminised Chinese face to expose the true emotions underneath. el Kaliouby and Colman (2020) write of the polite Chinese smile that ‘a naive observer might think this was a smile expressing happiness, but I knew better’. In doing so, she positions herself as an active and authoritative figure, ready to investigate and unmask the inscrutable Oriental. By positioning el Kaliouby as a kind of affective explorer who is able to crack the mask-like facade of Chinese people’s inner lives, *Girl Decoded* mimics the discovery narrative of the ‘voyeuristic Chinatown tour sub-genre of journalism’ of the late 19th and early 20th century (Yao, 2021: 194). The hidden figure of the Oriental women is thus unmasked by the software in a way that feels both physiognomic and voyeuristic in turn. Unlike autists, who are compelled to mask and reshape their affective expression and bodily habits to ascribe to Western societal norms of social relation, Chinese people must *be unmasked* by the liberal Western eye under Affectiva’s infopolitical regime.

Ultimately, Chinese people are portrayed as a counterpoint to the liberal, feeling individual who owns and experiences their own emotions; instead, emotions are collectively experienced and shared. Unlike the autistic case studies and examples, the discussion of Chinese customers and users is explicitly collective; Chinese people are only ever referred to through the plural ‘they’. This plays into a racial grammar that places Chinese people lower down in an affective hierarchy, and thus further from the liberal humanist ideal of the human. It compounds the long-standing racist trope that Chinese people are an indistinguishable collective, where individuals merge into an indistinguishable mass (Lester, 2021). In 1907, the famous British author Rudyard Kipling wrote that there were ‘three races who can work . . . but there is only one that can swarm’ (Lester, 2021: 1). This unending, alien sameness and the threat that it poses to White individuality is taken as a justification for the biopolitical exclusion – or even eradication – of Chinese people from Western societies (Lester, 2021: 2). While *Girl Decoded* does not make these genocidal claims, it evokes the same biopolitical logics that suggest there is interaction, but no individuality, among Chinese people. Chinese people, like autistic people, are considered devoid of liberal, individual personhood through their absence of the liberal subject’s capacity to normatively feel and express emotion. This fundamental lack is crucial for Affectiva’s ‘humanizing’ mission: for once el Kaliouby has affectively come of age, Affectiva must find new inhuman Others to teach and transform in the ultimate expression of infopolitical power.

Conclusion

In this article, we have examined what we term the *infopolitics of feeling* and interrogated how the power relations around feeling and unfeeling shape the technological and political project of facial ERT. Through our close analysis of el Kaliouby (and Affectiva's) coming-of-age story *Girl Decoded*, we have demonstrated that the ethical and political problems with ERT extend beyond the existing critiques of pseudoscience, poor performance, and privacy. Our reading of three key figures in el Kaliouby's autobiography – herself as the liberal feeling subject, the unfeeling autistic and the inscrutable Oriental – demonstrates how the capacity to feel and the ability to normatively express individuated emotion remain central to understanding what it means to be human, and to ascertain who is deserving (and undeserving) of residing under the banner of liberal humanity. Consequently, we call for a much wider interrogation of how the infopolitics of feeling shapes the field of ERT as a whole.

Acknowledgements

Many thanks to Dr Eleanor Drage for their thoughtful feedback on this research and for proofreading the final article. We would also like to thank Claire and Margaret Hopkins, and finally, each other, for this kind and thoughtful research process.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: One of the authors (Kerry McInerney) has been funded by Stiftung Mercator and has previously been funded by Christina Gaw and the Gates Cambridge Foundation.

ORCID iD

Kerry McInerney  <https://orcid.org/0000-0002-0053-4657>

Notes

1. Luke Stark and Jesse Hoey's (2021: 789) excellent analysis of the different theories of emotion and affect used and deployed by facial ERT systems contains a brief reference to Schuller's *Biopolitics of Feeling*, highlighting how 'emotional expression is a key vector through which racist hierarchies and misogynist tropes are produced (or "discovered") routinized, and enforced', while Kate Crawford (2021) articulates the taxonomies of feeling developed by Paul Ekman as a bridge between modern, computational analysis of affect and the racialised and gendered theories of the past. Similarly, Neda Atanasoski and Kalindi Vora's (2019) *Surrogate Humanity* draw on Schuller's work to examine the racialisation of social robotics and the programming of robot emotions. However, there is room for a much more in-depth analysis of how the capacity for emotional expression is used to taxonomise hierarchies and forms of life in the specific field of facial ERT.
2. We use the term 'Oriental inscrutability' not to suggest that the retrograde term 'Oriental' should be reintroduced into US popular discourse, but rather to connote a distinct racial history and set of stereotypes associated with Asiatic unfeeling.

References

- Adam A (2006) *Artificial Knowing: Gender and the Thinking Machine*. London: Routledge.
- Ahmed S (2014) *The Cultural Politics of Emotion* (2nd ed.). Edinburgh: Edinburgh University Press.

- AI Now (2018) AI now report 2018. Available at: <https://ainowinstitute.org/publication/ai-now-2018-report-2> (accessed 3 July 2023).
- Akash S (2022) Emotional AI market is expected to rise to US\$37.1 Billion by 2026. *Analytics Insight*, 2 February. Available at: <https://www.analyticsinsight.net/emotional-ai-market-is-expected-to-rise-to-us37-1-billion-by-2026/> (accessed 3 July 2023).
- Anderson B (2021) Affect and critique: A politics of boredom*. *Environment and Planning D: Society and Space* 39(2): 197–217.
- Atanasoski N and Vora K (2019) *Surrogate Humanity: Race, Robots, and the Politics of Technological Futures*. Durham, NC: Duke University Press.
- Baron-Cohen S (2002) The extreme male brain theory of autism. *Trends in Cognitive Sciences* 6(6): 248–254.
- Barrett LF, Adolphs R, Marsella S, et al. (2019) Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychol Sci Public Interest* 20(1): 1–68.
- Bastani N (2020) Education is the humanitarian's burden: Development and Iranian women's memoirs. *Cultural Critique* 109: 64–94.
- BBC News (2019) BBC 100 Women 2019: Who is on the list this year? Available at: <https://www.bbc.co.uk/news/world-50042279>
- Bennett CL and Rosner DK (2019) The promise of empathy. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*, Scotland, 4–9 May.
- Bow L (2022) *Racist Love: Asian Abstraction and the Pleasures of Fantasy*. Durham, NC: Duke University Press.
- Boyer A (2015) Data's work is never done. *Guernica*. Available at: <https://www.guernicamag.com/anne-boyer-datas-work-is-never-done/> (accessed 3 July 2023).
- Bui LT (2022) *Model Machines: A History of the Asian as Automaton*. Philadelphia, PA: Temple University Press.
- Cave S (2020) The problem with intelligence. In: *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Oxford, 19–21 May, pp. 29–35. New York: ACM.
- Cheney-Lippold J (2011) A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society* 28(6): 164–181.
- Cheng AA (2019) *Ornamentalism*. New York: Oxford University Press.
- Chun W (2009) Introduction: Race and/as technology; or, how to do things to race. *Camera Obscura* 24(1): 7–35.
- Code L (1991) *What Can She Know? Feminist Theory and the Construction of Knowledge*. Ithaca, NY: Cornell University Press.
- Crawford K (2021) *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University Press.
- D'Ignazio C and Klein LF (2020) *Data Feminism*. Cambridge, MA: MIT Press.
- Diprose R and Ziarek EP (2018) *Arendt, Natalty and Biopolitics: Toward Democratic Plurality and Reproductive Justice*. Edinburgh: Edinburgh University Press.
- Drage E and Mackereth K (2022) Does AI Debias Recruitment? Race, Gender, and AI's 'Eradication of Difference Between Groups'. *Philosophy and Technology* 35: 89.
- el Kaliouby R and Colman C (2020) *Girl Decoded: My Quest to Make Technology Emotionally Intelligent - and Change The Way We Interact Forever*. London: Penguin Business.
- Elam M (2022) Signs taken for wonders. *Daedalus* 151(2): 198–217.
- Gagen EA (2015) Governing emotions: Citizenship, neuroscience and the education of youth. *Transactions of the Institute of British Geographers* 40(1): 140–152.
- Gernsbacher MA and Yergeau M (2019) Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology* 7(1): 102.
- He H (2022) The racial interface: The computational origins of minority modeling. *Media-N*. 18(1): 31–52.

- Huang MN (2019) The Posthuman subject in/of Asian American literature. In: Rabinowitz P (ed.) *Oxford Research Encyclopedia of Literature*. Oxford: Oxford University Press.
- Huang VL (2022) *Surface Relations: Queer Forms of Asian American Inscrutability*. Durham, NC: Duke University Press.
- Kak A and West SM (2023) Confronting tech power. Available at: <https://ainowinstitute.org/2023-landscape> (accessed 3 July 2023).
- Keyes O and Austin J (2022) Feeling fixes: Mess and emotion in algorithmic audits. *Big Data & Society* 9(2): 20539517221113772.
- Kim E (2017) *Curative Violence: Rehabilitating Disability, Gender, and Sexuality in Modern Korea*. Durham, NC: Duke University Press.
- Koopman C (2018) Infopolitics, Biopolitics, Anatomopolitics. *Graduate Faculty Philosophy Journal* 39(1): 103–128.
- Lester Q (2021) Bio-orientalism and the yellow peril of yellow life. *Catalyst* 7(1): 34382.
- Lockhart JW (2020) ‘A Large and Longstanding Body’: Historical authority in the science of sex. In: Valencia-García LD (eds) *Far Right Revisionism and the End of History: Alt/Histories*. New York: Routledge, pp. 359–386.
- McStay A (2020) Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data & Society* 7(1): 4386.
- McStay A and Rosner G (2021) Emotional artificial intelligence in children’s toys and devices: Ethics, governance and practical remedies. *Big Data & Society* 8(1): 4877.
- Mantello P and Ho MT (2023) Emotional AI and the future of wellbeing in the post-pandemic workplace. *AI & Soc*. Available at: <https://doi.org/10.1007/s00146-023-01639-8>
- Murphy M (2019) *The Economization of Life*. Durham, NC: Duke University Press.
- Nagy J (2022) Autism and the making of emotion AI: Disability as resource for surveillance capitalism. *new media & society*. Epub ahead of print 15 July. DOI: 10.1177/14614448221109550.
- Nakamura L (2020) Feeling good about feeling bad: Virtuous virtual reality and the automation of racial empathy. *Journal of Visual Culture* 19(1): 47–64.
- O’Brien M (2021) Smart eye deputy CEO talks automotive AI. *Stuff*. Available at: <https://www.stuff.co.nz/motoring/300367854/smart-eye-deputy-ceo-talks-automotive-ai> (accessed 3 July 2023).
- Phillips A (2015) *The Politics of the Human*. Cambridge: Cambridge University Press.
- Picard R (2000) *Affective Computing*. Cambridge, MA: MIT Press.
- Podoletz L (2023) We have to talk about emotional AI and crime. *AI & Soc* 38: 1067–1082.
- Puig de la Bellacasa M (2017) *Matters of Care : Speculative Ethics in More than Human Worlds*. Minneapolis, MN: University of Minnesota Press.
- Roh DS, Huang B and Niu GA (2015) *Techno-orientalism: Imagining Asia in Speculative Fiction*. New Brunswick, NJ: Rutgers University Press.
- Rottenberg C (2018) *The Rise of Neoliberal Feminism*. New York: Oxford University Press.
- Schuller K (2018) *The Biopolitics of Feeling : Race, Sex, and Science in the Nineteenth Century*. Durham, NC: Duke University Press.
- Semel BM (2022) Listening like a computer: Attentional tensions and mechanized care in psychiatric digital phenotyping. *Science, Technology, & Human Values* 47(2): 266–290.
- Shah N (2019) ‘Asians are good at math’ is not a compliment: STEM success as a threat to personhood. *Harvard Educational Review* 89(4): 661–686.
- Sohn SH (2008) Introduction: Alien/Asian: Imagining the racialized future. *Melus* 33(4): 5–22.
- Srivastava S (2005) Tears, fears and careers: Anti-racism and emotion in social movement organizations. *Canadian Journal of Sociology/cahiers Canadiens De Sociologie* 31: 55–90.
- Stark L and Hoey J (2021) The ethics of emotion in artificial intelligence systems. In: *Facct’21: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*,

- 3–10 March, pp. 782–793. Available at: <https://doi-org.ezp.lib.cam.ac.uk/10.1145/3442188.3445939>
- Stark L and Hutson J (2022) Physiognomic artificial intelligence. *Fordham Intellectual Property, Media & Entertainment Law Journal* 32(2021): 922.
- Stevenson L (2014) *Life beside Itself: Imagining Care in the Canadian Arctic*. Oakland, CA: University of California Press.
- Su NM, Lazar A and Irani L (2021) Critical affects: Tech work emotions amidst the techlash. *Proceedings of the ACM on Human-computer Interaction* 5(CSCW1): 1–27.
- Ticktin L (2011) *Casualties of Care: Immigration and the Politics of Humanitarianism in France*. Oakland, CA: University of California Press.
- Tully J (2002) Political philosophy as a critical activity. *Political Theory* 30(4): 533–555.
- Van Dooren T (2014) Care. *Environmental Humanities* 5(1): 291–294.
- Wong D (2017) Dismembered Asian/American android parts in Ex Machina as ‘Inorganic’ critique. *Transformations* 29: 35–51.
- y Arcas BA, Mitchell M and Todorov A (2023) Physiognomy in the Age of AI. In: Browne J, Cave S, Drage E, et al. (eds) *Feminist AI: Critical Perspectives on Algorithms, Data and Intelligent Machines*. Oxford: Oxford University Press.
- Yao X (2021) *Disaffected: The Cultural Politics of Unfeeling in Nineteenth-century America*. Durham, NC: Duke University Press.
- Yergeau M (2013) Clinically significant disturbance: On theorists who theorize theory of mind. *Disability Studies Quarterly* 33(4): 231.

Author biographies

Dr Kerry McInerney (née Mackereth) is a Research Fellow at the Leverhulme Centre for the Future of Intelligence at the University of Cambridge, where she co-leads the Global Politics of AI project on how AI is affecting international relations. Her research explores the intersections between race, gender, technology and political violence through gender studies, politics and international relations, critical race theory, and Asian American/Asian diaspora studies. She is the co-editor of *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines* (2023, Oxford University Press) and *The Good Robot: Why Technology Needs Feminism* (2024, Bloomsbury Academic).

Os Keyes is a PhD Candidate at the University of Washington’s Department of Human-Centered Design & Engineering who researches issues of data, power and (in)justice. An inaugural Ada Lovelace Fellow, their previous work has included studying questions of gender, disability, race and power in biometric security systems, the dynamics of power in the viability of “algorithmic audits,” and injustices in the discursive shaping of labour practices around Artificial Intelligence.