Opacity, Transparency, and the Ethics of Affective Computing

Manohar Kumar, Aisha Aijaz, Omkar Chattar, Jainendra Shukla, *Member, IEEE*, and Raghava Mutharaju

***Abstract*—Human opacity is the intrinsic quality of unknowability of human beings with respect to machines. The descriptive relationship between humans and machines, which captures how much information one can gather about the other, can be explicated using an opacity-transparency relationship. This relationship allows us to describe and normatively evaluate a spectrum of opacity where humans and machines may be either opaque or transparent. In this paper, we argue that the advent of Affective Computing (AC) has begun to shift the ideal position of humans on this spectrum towards greater transparency, while much of this technology is shifting towards opacity. We explore the implications of this shift with regard to the affective information of humans and how the threat to human opacity by AC systems has various adverse repercussions, such as infringement of one's autonomy, deception, manipulation, and increased anxiety. There are also distributive consequences that expose vulnerable groups to unjustified burdens and reduce them to mere profiles. We further provide an assessment of current AC technology, which follows the descriptive relationship between humans and machines from the lens of opacity and transparency. Finally, we foresee and address three possible objections to our claims. These are the beneficence of AC systems, their relation to privacy, and their restrictive capacity to capture human affects. Through these arguments, the paper aims to bring attention to the ontological relationship between humans and machines from the perspective of opacity and transparency while emphasizing on the gravity of the ethical concerns raised by their threat to human opacity.**

***Index Terms*—Affective computing, opacity, transparency, ethics, autonomy, distributive justice, manipulation.**

I. Introduction

RECENT advances in affective computing (AC) have changed the way we operate and interact with our surroundings. The very intrinsic features of our being, such as our emotions, moods, tone of voice, facial expressions, and way of walking, aspects we have little control over, can now be captured through AC devices. There are various beneficial applications of such systems. Wearable devices that can measure heart rate can predict if and when someone is at risk of a stroke or heart attack [1]. Individuals diagnosed with autism may also use AC systems to better express their emotions [2]. It allows 'the user to feel as if his or her strong affective state has been effectively communicated' [3]. Early intervention systems based on affective technology aid people with developmental disorders to live a more fulfilled life [4]. Artificial companions and support agents are also used to care for the elderly [3], making the lives of individuals and their families easier [5]. AC systems like MACH (My Automated Conversation Coach), an interview preparation tool, is used to help candidates prepare beforehand so they may be more confident in scoring their dream job [6]. Thus, AC has a significant impact on human lives and aids people in making decisions, taking action, and interacting intelligently with both machines and their peers [7].

Notwithstanding its positive impacts, the deployment of AC systems raises serious ethical [8] and meta-ethical concerns [9] like privacy [10], deception and emotional dependence [3], [11], calling for design contractualism [12], [13]. These ethical problems have been majorly understood from the design and deployment perspective. However, an ethical assessment of AC systems requires understanding the consequences that arise from the very constitution and structure of AC, which alters the relationship between humans and machines. In this paper, we argue that this relationship should be understood within a spectrum of opacity and transparency. Normative discussion frames transparency as a demand from machines while leaving opacity for humans. AC, however, relies on shifting this fundamental relationship, where human opacity is directly impacted. This paper investigates the ethical implications of AC on human opacity and offers a possible way to balance competing considerations of opacity and transparency.

This paper is structured as follows. First, we outline an opacity-transparency relationship that structures the relationship between humans and machines. We argue that this relationship is both normative and descriptive that frames the ideal relationship between humans and machines. We then investigate the implications of this view by examining the ethical implications of AC by examining it on the opacity-transparency spectrum. Finally, we examine some objections to our view.

Manuscript received 6 June 2022; revised 7 April 2023; accepted 3 May 2023. Date of publication 23 May 2023; date of current version 1 March 2024. Aisha Aijaz and Omkar Chattar are supported by IHUB ANUBHUTI-IIITD FOUNDATION Doctoral Fellowship (Reference ID. IHUB Anubhuti/Project Grant/04. We acknowledge the support of the IHUB-ANUBHUTI-IIITD FOUNDATION set up under the NM-ICPS scheme of the Department of Science and Technology, India. In addition, Manohar Kumar, Jainendra Shukla and Raghava Mutharaju are partially supported by the Infosys Centre for Artificial Intelligence, IIIT-Delhi. Further, Jainendra Shukla is also supported by Center for Design and New Media (Sponsored by Tata Consultancy Services, A TCS Foundation Initiative). Recommended for acceptance by Special Issue on Ethics in Affective Computing. *(Corresponding author: Aisha Aijaz.)*

The authors are with the Indraprastha Institute of Information Technology, New Delhi, Delhi 110020, India (e-mail: manohar.kumar@iiitd.ac.in; aishaa@iiitd.ac.in; omkarch@iiitd.ac.in; jainendra@iiitd.ac.in; raghava.mutharaju@iiitd.ac.in).

Digital Object Identifier 10.1109/TAFFC.2023.3278230

II. The Opacity Transparency Relationship

The opacity-transparency relationship is a descriptive relationship between a human and a machine that captures the extent to which either entity can know about the other and the extent to which one's internal information is opaque to the other. Knowing in the case of machines would be knowing how they function, and in the case of humans would be knowing their intrinsic information such as their emotions, biases, preferences, etc. We begin by clarifying the terms opacity and transparency in order to describe the nature of their relationship.

A. Opacity

The Oxford Dictionary defines opacity as a quality of an object or a person that demonstrates 'the fact of being difficult to see through'. In other words, opacity is the intrinsic quality of unknowability of an object that makes it difficult to gather information. One of the primary characteristics of opacity is the resistance to legibility. In their anthropological work on opacity, N. Buitron and Steinmuller argue that legibility refers to making something decipherable and interpretable [14]. It enables distinctiveness for the observer to categorize the observed entity or surrounding, thus forming a relationship between them [14]. A human being is opaque when features intrinsic to them are not available for scrutiny.

On the other hand, machines are considered opaque when they render their functioning unknowable to human. For Introna, an opaque technology has uses that are obscure and passive functionality that allows limited operation by the user [15]. It is difficult to trace the operation from the instruction to the decision process in such systems.

B. Transparency

Transparency, on the other hand, entails knowability and scrutiny in terms of humans. A transparent technology has its operation on the surface, is in use with fair guidelines, has application stability, and is transparent in its outcome and use [15]. Transparent technology is both phenomenologically and procedurally transparent. Wheeler, following Heidegger, argues that technology is phenomenologically transparent when it enables skillful manipulation in'a hitch-free manner' without any conscious acknowledgment of the item in use [16]. Procedural transparency, or what Wheeler calls an "open to understanding" conception of transparency, allows a class of users to understand how the machine operates [17]. For our purposes, the kind of transparency that we are interested in is the procedural kind.

C. Opacity Spectrum

The opacity-transparency relationship is not dichotomous as an either/or relation [15] but rather exists on a spectrum with varied degrees of transparency and opacity of the machine and the human. We call this the opacity spectrum. It is inspired by Bert Koops et al. privacy spectrum and Introna's idea of a transparency continuum [18] [15]. For each set of technology, we can visualize the transparency-opacity relationship on a spectrum that captures continuity, shifts, and various tensions on transparency and opacity between humans and machines.

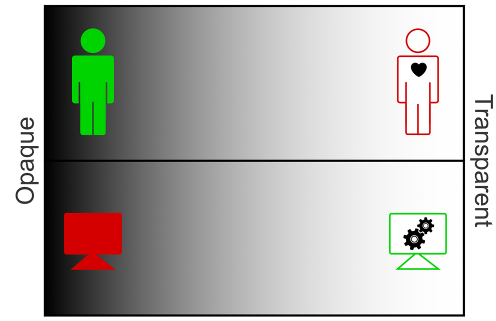


Fig. 1. A visualization of the opacity spectrum. An ideal normative relationship between humans and machines would imagine retaining human opacity (Green for the human) and complete transparency of the machine (Green for the machine). While the red represents a scenario where the machine is completely opaque and the human is completely transparent. The spectrum represents the move from a greater to reduced opacity for human beings, and greater to reduced transparency for machines. Actual world use cases will lie at different points of the spectrum.

Technology such as table clocks, x-ray machines, and calculators are procedurally transparent in the sense that we know how they operate and the extent of their function. However, smartwatches and phones are not procedurally transparent about their functioning and are thus opaque. To some degree, both developers and users of such technology do not know every aspect of their operation. Developers cannot pinpoint how or why such black box systems make a decision, and users do not know the extent of their function, especially when they are connected to a network and have the ability to discreetly transfer information. This functioning of contemporary systems can be best captured on a continuum. On one end of the spectrum, we place artifacts whose functioning is transparent to us and we know what consequences their operation has for human beings. On the other end are artifacts that give us little to no information about their functioning. Fig. 1 represents the relationship on a spectrum.

Although Invasive technologies shift the position of humans on the opacity spectrum towards transparency, not all sit close to being completely opaque themselves. For example, a simple microphone that captures voice is procedurally transparent to us. However, given that the current trend for most technology is a shift towards autonomously learning models, most contemporary systems function on such in-the-dark algorithms. We are thus inclined to consider these, particularly as we experience greater intrusiveness with increasing human transparency and opaque systems. This is in contrast to traditional technical tools that are highly transparent while human beings are opaque.

Thus, the opacity spectrum enables us to capture the loss or enhancement of transparency from innovation and emphasizes the fuzzy nature of opacity loss. The degree of transparency varies depending upon whether the data gathering is localized to generate use value or is stored in servers and combined with other data to create a profile, leading to even greater exposure of human beings.

D. Opacity-Transparency as a Regulative Ideal

The transparency-opacity relation in phenomenological and post-phenomenological approaches is an ontological relation [19]. As technology advances, we note that the ontological relationship shifts towards reduced transparency, manipulability [20][1], and understanding of machines which undermines the capacity to control them. This is because an opaque machine forces humans into transparency by gathering their personal information, putting them into profiles, and making predictions without their explicit knowledge or understanding. The shift in the ontological relation then enables a greater degree of control about the knowable object/subject. This reduces the subject's control and compromises the autonomy that comes with their opacity.

Glissant offers an ontological and political account of opacity, especially in a language that rests on the fundamental reduction of humans into existing and legible categories of knowing [21]. Opacity is a condition that is integral to the human being that resists, according to Birchall, the demand to be "knowable, understood, measured, categorized, and rendered transparent" by the dominant political order [22]. Opacity thus is "an ontological condition of minoritarian subjectivity and r(el)ationality" [23] and is thus "both a condition of and a relation of freedom" [21] that resists assimilation and objectification.

This analysis has specific relevance for invasive technologies that reduce opacity. As Brey argues in his analysis of facial recognition, they turn the face, intrinsic to a human and with immense personal value, into an instrument, information, and ultimately data [24]. This functional reduction creates data about a person's body outside of their control. This reduces autonomy over one's own body and is particularly relevant to vulnerable groups or minorities who have no way of understanding the mechanisms of decision-making. This change follows the integration of technical dependency in everyday life, where humans are subjected to an empirical investigation resulting in their objectification, manipulation, and infringement of autonomy. The shifting opacity-transparency relationship then does not merely remain a technological phenomenon but rather a psycho-phenomenological one as well.

Enactments of opacity in digital environments argues Birchall, enable subjects who are "doubly afflicted by a lack of opacity–ontological and digital" to "interrupt the violence of transparency inherent in power" and enable us "to begin to rethink the role of sharing in a data ecology that demands visible, surveillable, quantifiable, and entrepreneurial subjects" [23]. Opacity enables individuals to maintain intentional control over information through secrecy [25] or with deliberate obscurity that affords them a sense of protection [26], [27]. Opacity supports plurality by eliminating the need for collective choice or an official public stance [28] and is thus fundamental to the formation of identity [25]. It allows individuals to regulate, control, and make decisions about their emotions, who to share their information with, and consequently, build relationships of trust. Forced transparency increases dishonesty and deception, reduces trust, makes people evasive, and encourages half-truths, hypocrisies, and forms of 'self-censorship' [29].

The centrality of opacity to human autonomy does not undermine the role of transparency in fulfilling significant human needs. As Zach Blas argues, transparency is important for the vulnerable when it enables them to fulfill some need like the desire to be recognized [30]. Similarly, diagnostic technological tools like X-rays fulfill the human interest in better healthcare by enabling individuals to know whether they are at risk of a life-threatening disease. Another example is that of surveillance technologies that aim to capture crime and reduce the threat to human safety. Human beings have fundamental interests in their safety, and to that extent, these technologies are justified. But these technologies may also impinge on privacy and threaten civil liberties. They raise distributive concerns of justice regarding how the benefits and burdens of such security policies are distributed [31], [32] especially post 9/11, where invasive technologies have been deployed to tackle threats to security. Thus, concerns for opacity and transparency follow a greater demand for transparency of invasive technologies [33]. This requires explainability, justification, transparency, and accountability of decision-making [34] and requires that institutions that make decisions on behalf of the individual be able to explain, defend, and justify those decisions.

The opacity-transparency relationship is thus a normative tool for what Brey, in his work on disclosive ethics, calls 'moral deciphering of computer technology' [35]. He explains that disclosive ethics relates to the disclosure of moral features and evaluation of normativity with regard to computer technology. The two-stage process includes the analysis of the technology from a moral perspective and the development of moral theory, which is later refined. This method is sensitive to the changing circumstances where new technologies require an assessment of older values and the development of new ones. Conceptualizing opacity-transparency relationships on a spectrum enables us to then assess new developments in technology from the perspectives of opacity and transparency, values that protect human autonomy and are key to concerns of justice. If a new technology threatens fundamental human opacity, autonomy, or other human interests, it may suggest reasonable grounds to reject or modify it unless it comes with a substantive benefit for human beings in terms of protecting substantive values. In other words, this relationship can also frame the *normative* opacity-transparency relationship by encouraging an adequate balance.

III. Ethics of Affective Systems

Affective computing, in the words of Rosalind Picard, is about "building machines that have several affective abilities, especially: recognizing, expressing, modeling, communicating, and responding to emotion" [7]. The promise of affective computing lies in the capacity of computers "to recognize its user's affective expressions and to respond intelligently, especially if the user

[1] Turkle argues that early users of computers demanded transparency from the machine where they could look in the machine and understand how it functioned, a reason why early Macintosh computers were criticized for their opacity. The Apple system had introduced a conversational interface with no hint of its inner workings, thus merely displaying an 'artful navigation of opacity' [20].

indicates frustration, fear, or dislike of something the computer can change" [36]. Understanding what frustrates users enables designers to make improvements in the computer tools [36]. This is especially important because most people are unaware of how computers "work and find computers and software increasingly hard to understand". Users are thus "caught in a spiral of increasing complexity" that leads people to change what they expect from the system. These systems, she argues, are far from human-centered, where they are not even aware that they have upset 'their most valuable customer'. Affective computing, on the other hand, aims to make computer systems more acceptable and naturalistic by being sensitive and responsive to a human's affective state [37]. It assumes that humans demonstrate patterns that have physical giveaways such as fluctuating heart rate, pupil dilation, and muscle tension [38] and convert these effects into observable data.

Affective systems, it is argued, will enable a better quality of life [3], [7], aid human beings in making rational decisions [39], problem-solving and creativity [40], enable skillful listening [41], reduce user frustration [42], and improve emotional and experiential skills [7]. Additionally, the benefits of AC in healthcare, such as preventing seizures and aiding children with autism spectrum disorder, are also significant. AC can thus enable cognitive and affective enhancement similar to technology such as a diary and a calculator [43]. The way a diary is an essential tool for someone with Alzheimer's [44], AC is an affective enhancement that enables human function, especially for those with a disability. In this regard, AC enhances autonomy and opens up space to exercise their capacity as well-endowed beings. However, the potential for affect recognition also opens up ethical questions primarily due to its epistemic foundations and its deployment in decision settings. It is to these aspects that we turn next.

A. *Epistemic Concerns With AC*

An AC system discerns and classifies emotions into certain categories. In AC systems, Boehner et al. argue, "affect is often taken to be another kind of information–discrete units or states internal to an individual that can be transmitted in a loss-free manner from people to computational systems and back."[45] This process is a form of reduction of emotion into data based on a classificatory process and violates the person's opacity as it captures internal information about them. However, this process rests on assumptions about the nature of emotions, the criteria used to classify them, and the allocation of these into categories. This raises crucial epistemic questions regarding the interpretability of emotions, their adequate classification, and the possibilities or not of misrecognition, misidentification, and misinterpretation. They further raise questions about the credibility of the machines' claim to know these emotions and their evaluation against competing claims by humans.

Emotions are amongst the least understood of human experiences [46][47]. They have evaluative, physiological, phenomenological, expressive, behavioral, and mental components [48]. It is unclear which of these components is most essential to emotions. Disagreement continues whether affects depend on the physicality of the human body or on our context and social background [49]. An emotion that represents an act of communication [50] has significant implications on the message received. Individuals have varied motivations to express their emotions that may or may not represent their mental state. An individual demonstrating embarrassment may actually be doing it on a social cue rather than actually feeling it [51]. Perceiving emotions in a natural environment is much more difficult than in a specifically designed one [52]. It is more than just applying categories to complex emotions and requires awareness of the state of the person, the space they are in, and other cultural factors [53]. Some research does focus on the integral role of facial expressions in the demonstration of emotions. However, there is a counterargument that requires multi-modal approaches registering other affective states to study complex human emotions [54]. This may include the tensing of muscles, bodily movement and gestures, perspiration, and even physiological affects such as heart rate and pupil dilation. Thus emotional perception would need recognition of context, cues, and individual background apart from direct emotional and physiological effects. This makes one think of the many opaque layers an AC system would have to invade in order to make an accurate decision about the emotional state of a person.

Deciding on classes to categorize the emotions that stem from these many factors is thus a challenging task [52]. This was also discussed by Cowie, who recognized the importance of understanding culture and nature apart from other cues such as appraisal, impulses, actions, reactions, involuntary signs, and the source of the emotion [55]. Context also plays an important role in emotion detection [56]. Context does not merely include the circumstance or the interaction that induces the affect but also the socio-cultural and political context. For example, a wave of the hand may mean either 'hello' or 'no', depending on cultural context. Similarly, a frown may be interpreted as "anger, anxiety, disgust, and contempt, when a person is simply deep in thought" [56], [57]. Therefore, the assumptions of the universality of emotions are shaky [58].

At the core of AC lies the problem of accurate definition and classification [59]. Affect capture is normative judgment based on a theory that affects are natural and universal that "posit universally applicable emotion taxonomies and a fixed relationship between people's inner states and their behavioral displays"[60]. This, Luke Stark and Jesse Hoey argue, "can emerge from conceptual assumptions, themselves grounded in a particular interpretation of empirical data or the choice of what data is serving as proxy for emotive expression." [61]. Affects here are understood as information, a 'dual of cognition' which 'operates in concert with and in the context of traditional cognitive behavior'. Due to this conception, emotion perception suffers the "difficulty in accounting for and adequately incorporating an understanding of everyday action as situated in social and cultural contexts that give them meaning" [45]. An element of partiality is thus introduced, reducing the extent of emotion capture and affecting its judgment and actions. As Lisa Feldmann Barrett argues that even if one were to assume that affects are natural categories, there is no way of finding out because it is produced in response to a contingent circumstance and is thus not reducible to a pre-existing category [62]. Even earlier understandings of emotion, such as those premised on a face-coding model presented by Paul Ekman [63], now stand disputed [59]. Emotions cannot be treated as mere static entities. Rather, they should be considered as dynamic, ever-changing aspects of the human experience [64], [65]. This challenges the way emotion datasets are organized [66]. The reduction of emotional states to relatively basic categories may have some beneficial consequences for limited tasks. However, it fails when the challenges are more complex.

In order to understand emotions and make meanings from them, machines require disambiguation of the speakers' utterances and appeal to relevance, intent, and context [67]. Allowing a machine to judge our emotions gives them space for interpretation and action. The way the dataset is interpreted will have significant implications for its decisions and those affected by them. If an AC system is unable to read the problem correctly, it will offer the wrong solution. Its epistemic state will be compromised in the process, compromising the standing of the person vis-a-vis the system. Thus, AC systems potentially hold possibilities of misrecognition and wrong characterization [68], which will have crucial implications the more these systems are integrated into society and in crucial decisional spaces. In a situation where the outcome of a decision made by an AC system is to be judged, the lack of its transparency would lead to a lack of accountability as well.

There have been claims regarding advanced computer systems' ability to know the individual better than themselves [69], [70], and their ability to capture affect, sexuality, and personality [71]. This already alters the frame through which claims on emotions are assessed. However, in circumstances where the individual is unable to adequately describe or make sense of their own emotional experiences, perhaps lacking the lexical repertoire or language required to express themselves [72], AC may enable a form of interpretability that allows them to overcome a feeling of frustration. It is already proving useful for autistic individuals where behavioral interventions are required to aid communication difficulties [2]. But when our affects are understood, mediated, interpreted, and predicted by a machine, it changes the way we make sense of them. Human beings may begin to understand their emotions more through the mediation process. Cognitive enhancing devices like turn-by-turn GPS navigation apps make us surrender the critical skill of spatial navigation to machines [73]. Mediation as a form of enhancement does not pose moral questions per se, as humans seek enhancement that enables their functioning. However, it becomes problematic when it becomes the hermeneutic universe through which human emotions are judged and categorized. These systems may exert a normative influence by guiding emotions to suit the particularities of the context and conformity with a norm, a profile, thus presenting sites of conflict [53] that may require users to adapt their natural, intrinsic emotions according to the needs of an artificial machine [74]. For example, in cases where interviewees are aware of an AC tracking their emotions, they tend to game the system based on its expectations [75] and start behaving in unnatural ways, like machines themselves [76].

In interviewing systems, candidates have demonstrated a ”tendency to glorify the technology,” and this ”made them trust it would make better decisions than human ones. This resulted in them feeling 'judged' by a sort of superior entity.,” [76] and were thus ready to please the technology [77], This demonstrates that candidates assign greater credibility to the system, a consequence also of the credibility that the hiring platforms ascribed to the algorithm. The three hiring platforms analyzed by Zahira Jaser et al.'elevated the selective, unbiased power of their technology' and 'overplayed the validity and reliability of the results so much so that candidates believed that their non-selection is more a result of their own lacking than that of the biases of the technology [77]. Albeit Franziska Hirt et al. did not find any connection between the estimate of emotions by ACs and a subjective assessment of the emotions based on a self-report [78]. However, it is not implausible to imagine that once AC becomes ubiquitous, and is introduced in the marketplace, the way we assign credibility to it may change [60] as is evidenced in the widespread deployment of facial recognition technology despite doubts regarding their efficacy.

This claim will be clear with another example. Affective systems are increasingly being deployed for workplace surveillance [79]. Emotion surveillance is considered important in settings where workers are required to demonstrate pro-social attitudes, which is cited as a reason for failure to achieve targets. [80] Despite the claims made by workers that the failure owed to systemic reasons of inadequate data, slow systems, or inaccurate data, they were forced to go through a 'performance counseling scheme'. This created a kind of exposure that made the employees emotionally vulnerable and anxious. The assessment by technology reduces a social interaction into an adjudication of emotions as being either negative or positive. Workplace surveillance (or what Kirstie Ball calls Exposure) creates datafied subjects and a credibility structure where the worker is rendered suspect and thus responsible for failure. It extends greater credibility to the machine than the person themselves. This pushes the burden of justification on humans. In this regard, they experience a double disadvantage that of greater exposure and control and a reduced value to their testimony. Assigning greater testimonial credibility to machines raises questions about the epistemic validity of the human assessment, thus doubting their claims as primary knowers. This represents an instance of testimonial injustice that is a genus of epistemic injustice [81] where the human claim as a knower is accorded less primacy than the claim of the machine. This is especially important because such forms of exposure are both opaque and have an institutional element to it. Exposure of this kind changes the opacity-transparency relationship towards lesser opacity of the human and greater opacity of the institutions (if not the machines) that deploy them. In this regard, this deployment requires that greater discretion and increased transparency are demonstrated, and these systems should only be deployed if only essential because it reduces the autonomy of the individual, subjecting them to manipulation and control. This we discuss in the next section.

B. Affective Computing and Concerns of Autonomy

Emotions play a complex function for humans. Individuals are always in the process of representing themselves and act differently in response to their environment, the context, and how they interpret their role in the situation [82]. They make conscious decisions to reveal certain aspects of their lives as part of this representation and retain others [83]. Affects are anthropomorphic happenings that aid us in this representation and enable us to retain the capacity to be opaque in terms of our intentions, motivations, and emotions. Being unpredictable is linked to human agency and allows individuals to appear as they want to [84]. Humans are also unable to predict what kind of self their actions reveal and how it will be interpreted and further communicated by those in their vicinity [85].

Unwarranted profiling of affects leads people to act in ways they are neither predisposed to nor are in their interest [86] and alters our interaction around machines [53], thus undermining their right to express their emotions and have an emotional culture fostered on them [87]. Therefore, ACs unable to capture and appreciate differences, integrate idiosyncrasies, and recognize contextual responses of individuals, effectively undermine the communicative potential of an individual's emotions, disciplines human gestures, and undermine the demand that human emotions be understood on their own terms. This compromises the hermeneutic universe in which humans make sense of their emotions and forces us into an affective surrender we perhaps still do not comprehend. Candidates exposed to affective systems in interview settings have claimed to feel diminished in their humanity and depersonalized feeling that they were behaving like robots [77].

Emotion profiling also undermines the possibility of self-definition; an individual may not be able to experience the tension between nascent feelings (such as those leading to anger) and one's meta-preferences or aspirations (for instance, the kind of parent one strives to be). Instead, the individual is presented with an instantly 'optimized' environment that either removes or masks the cause of one's upcoming anger. This optimization may enable efficient decision-making and smoother interactions but may also leave us with a poorer version of ourselves, one that is deprived of the chance to learn from and grow through the experience of such tensions. This has perilous consequences by exposing individuals to deception and manipulation that we turn to next.

1) Deception: AC systems capable of displaying emotions to form bonds with human beings are quite common [88]. Matthias Scheutz argues that socially assistive robots demonstrating affects demonstrate care for the human subjects and may generate positive responses in those cared for [89]. Despite its beneficial aspects, its implications for autonomy are particularly worrisome. These robots may give a false impression that the affects they demonstrate are actually internal states, i.e., the affects are not merely displayed but are also felt and are thus enough for people to develop emotional dependencies and relationships with them. Humans require that ACs get their affects right and respond to the demands of context appropriately. If they fail to do so, they can cause them suffering [89]. Although these systems have been found to aid in learning, they can generate expectations in their users that they are in no position to fulfill [90]. Children, for instance, can develop bonds, social connections, and entanglements with robots [91]. These developments are particularly concerning for teachers who believe this may raise questions of trust and deception [92]. This constitutes a kind of deception that may lead to stress and false trust [93]. This deception may even occur with robots that are not designed to deceive and where human beings are aware of such possibilities [88].

Secondly, care robots raise a crucial concern regarding the expectations they generate. Humans do not want robots to get their affects right for them to feel cared for, but they actually want to be cared for. Caring requires more than a presence of affects. It demands "attentiveness, responsibility, competence, responsiveness, and trust (and solidarity)" [94]. It entails a commitment that is both emotional, reflective [95], and dialogic in order to refine the care based on need [96]. Care is not a pre-reflective ethical commitment but requires connection and awareness of the context, [97], requires listening and sentimental communication [98]. It requires assuming responsibility and the recognition of the need for care [99]. It may be argued that robots basically simulate the expressions of doctors and nurses who demonstrate an ethics of care without necessarily getting affected by them [100]. This reduces those cared for to an instrumental dependence on receiving a service that entails getting better but does not provide the emotional bond that is crucial for care.

2) Manipulation and Reduced Self-Determination: Manipulation is "a form of influence that is neither coercion nor rational persuasion" [101]. It bypasses rational deliberation, is deceptive in nature [102], induces false beliefs and expectations [103], and is a form of hidden influence [104] that leaves us with lesser control, affects our self-respect [105], and reduces human beings to things [101].

Manipulation of affects is quite common on online platforms [106], [107] that through 'techno-social engineering' and profiling influence how we think and act [108], [109]. Transparency of emotions enables machines to influence our decisions, actions, and judgment [110], [111]. In this regard, ACs engender the constant possibility of being watched, heard, and kept track of, thus exposing us to the threat of manipulation. In June 2014, Facebook published a study stating that it was able to manipulate emotions in its case subjects by means of emotional contagion that the subjects were not aware of. The study had a large sample size of 689,003 [112]. Studies show that Facebook has been using emotional manipulation to actively keep users engaged with the platform in a way that only benefits corporations [113].

Potential advertisers have a keen interest in the affect data, [114] because they believe that if 'one can sway emotions, one has a better chance of influencing cognition'. This is despite their awareness that this is manipulative. The collection of affective data creates vulnerabilities when user profiling is used to generate more revenue [115] by encouraging 'sad-spending' [116]. Similarly, it may lead companies to have access to mental health data that may not be protected under the Health Insurance Portability and Accountability Act of 1996 (HIPAA) [117]. This may lead to third-party users misusing such data to push for more aggressive advertisements, planned campaigns, spreading deceitful information to vulnerable populations, and taking advantage of their emotional and mental vulnerabilities [118]. Recent revelations by a Facebook whistleblower, Frances Haugen, made it public that the social media giant refused to take down hateful content and actively promoted it as it led to an increased engagement of users [119]. Targeted advertising has also been used to influence user behavior, especially in political campaigns and social decisions [120], [121] [122].

AC-related manipulation raises moral concerns at two levels. At the first level, where the AC device is necessary for the functioning of the human being, any attempt to alter or manipulate its functioning will affect them and their environment. For example, those at high risk of seizures are dependent on the AC system, and so are children with autism. For these persons, the AC system is irreplaceable [123]. Any interference with their functioning not only affects how they make sense of their environment and orient themselves to the world but also makes them vulnerable. This is similar to cognitive enhancement in patients with Alzheimer's where interference with their cognitive artifacts raises moral concerns [43], [124]. In a way, emotional manipulation undermines the mental integrity of the person [111].

The second aspect is that humans with access to these technologies can exercise power over the ones who do not. Humans thus stop being ends in themselves and rather become means to those who have access to these technologies who can then exercise power over them. These forms of manipulation thus not only reduce our individual autonomy but also affect our self-determination as members of a political community. This has perilous consequences for democracy in the long run.

C. Distributive Effects of Affect Recognition

1) Reduced Trust and Enhanced Anxiety: The use of ACs has far outstripped its original mandate of smoothening human-computer interface. They have been deployed for the purposes of worker surveillance to gauge the dispositional attitudes of workers and, in more extreme cases, to keep track of how people behave and respond to certain political or social propaganda [125]. These are not merely altruistic innovations that enable human capacities and reduce frustrations but rather increase the stress and anxiety of those exposed to them. For instance, AC, as Lachlan Urquhart et al. argue, is altering the workplace by 'enabling technologically mediated professional relationships.'[126] This may have benefits 'like protecting worker well-being' when used wisely but can also engender significant risks 'such as forms of tracking that benefit employers at the expense of employee interests'. The deployment of ACs in the workplace though is rarely about workers' welfare but is driven by considerations of efficiency and profits. Affective workplace surveillance targets the workers' 'thoughts, feelings, and physiology' [80] the 'hitherto 'unmarked' or 'un-inscribed' aspects of the subject", as it 'surfaces and becomes enacted at the body boundary [127]". Institutions that do surveillance engage in an economy of interiority that can indirectly affect autonomy by enhancing the stress and anxiety of the employees through the process of granular tracking and gamification [128], [129]. This forced transparency makes the employee vulnerable to control and manipulation [130][131], control over their time and bodies of the workers and exposes them to various kinds of injuries and vulnerabilities [132], and undermines the trust relations between workers and supervisors [133] [132]. This lowers the capacity of the employee to maneuver and operate in their work environment, with the assurance that their actions and utterances do not work against them.

Workplace surveillance exposes workers to the ire of their management by failing to account for context and 'human particularities such as attitudinal diversity, gender differences or cultural idiosyncrasies' [49], but are rather predictive models that assume statistical continuity based on past evidence of productivity. In situations where an employee is not feeling well, having a bad day, or experiencing problems at home, the context is lost to the system, which simply profiles based on current perceivable affects. Not only is this wrong, but it diminishes sympathy and interpersonal conversation between the management and their employees. Thus, affective workplace surveillance raises distributive concerns where the burdens of surveillance activities are disproportionately borne by the workers at the cost of their emotional and physical well-being.

2) Affect Recognition and Distribution of Resources: Affective computing deployed in distributive settings raises concerns of distributive justice [134]. Whether affect detection leads to an equal distribution of benefits and burdens or affects marginalized groups disproportionately has implications for its justification. So wherever affect detection plays an important role in making decisions, individuals can demand that their affects are well understood in ways that they would present them and not manipulated, misrepresented or misread.

Consider the deployment of AC systems to track the demonstrated interest of college applicants captured through their interaction with various links on the university website and the amount of time they spend on it. Demonstrated interest is seen as important to decide between applicants at the borderline [135]. However, it does not account for the mixed motivations of applicants and takes away the possibility of reconsideration, thinking it over, weighing their options when confused, and deciding what is best without being denied the opportunity of what may be a second-best option. Tracking demonstrated interests thus limits the applicants' choices and demands fidelity in set-ups where it is not required; one is not required to demonstrate fidelity to a college from the very beginning.

Distributive consequences may also arise through profiling that confers disadvantages or advantages based on the profile. Profiling is a form of classification that categorizes individuals into groups based on some similarity of attributes and then targets them similarly. Emotional profiling may be used to detect suicidal individuals on social media [136], to reduce threats of accidents on roads [137], for recognizing signs of depression [138], documenting suicidal thoughts [139], identifying stress in individuals through wearable devices [140]. Treating individuals based on a shared profile is not wrong [141], [142], [143]. However, the lack of differentiation between individuals within the same profile may create bias in distributive decisions [144], [145]. Where individual differences hold moral weight, justice requires that those differences do not put those exhibiting them at a disadvantage. This is where profiling, especially emotional profiling, raises ethical concerns. In the context of AC, we need to consider (a) whether the profiling is done accurately and (b) whether the user is aware of such a classification [146].

Instances of (a) have been raised in the reliance on facial and affect recognition to make decisions in interview settings [147]. This, companies claim, enables faster hiring by algorithms trained on reliable data from successful interviews, where hired candidates went on to become true assets to the employer. Hirevue, an interview tool used by companies, employs algorithms that learn from a 25000 database of previous successful interviews, including affects such as voice tone modulation, talking speed, raised chin, smiling, raised brows, wide or closed eyes, and tightened lips [147]. This system assumes a uniform distribution of attributes while only being trained on attributes of mostly white men. Cues such as confidence and self-assurance in successful applicants are a product of social construct and are distributed differently among various genders, races, ethnicities, and religions [148]. Algorithms trained on attributes of one group of people perpetuate already existing inequality. Profiling here is unjust because it goes against fundamental interests, such as qualifying an interview and undermining an individual's potential to communicate and represent themselves, further reducing their sense of self and leading to social cruelty [85].

This leads us to point (b). Profiling operates in a space of epistemic asymmetry. There is a systemic lack of oversight [149] from the black-box approach that effectively hides information [59] and renders the user unaware of the information being stored, the extent to which it may be used [144], and why they are being understood in a certain way. Candidates have actually demonstrated ignorance when interviewed by AC systems about the role that the algorithm plays in the interview process. Companies deploying them failed to offer a clear understanding of the functioning and the limitations of the technology. So when affective systems are deployed in decision-making settings, and a candidate faces a rejection, it is unclear who is ultimately responsible for making the decision. In the absence of a proper accountability and responsibility framework, there is a gap that impinges on the autonomy of the individual to be provided with answers for decisions that affect them negatively. The consequences of affect detection and the absence of accountability structures on autonomy are particularly worrisome in the case where emotions displayed by machines open up spaces for deception and manipulation.

IV. Discussion

A. *Opacity-Transparency in AC*

So far we have demonstrated the epistemic and ethical considerations of AC. We noticed that AC offers various benefits in education, healthcare, security, and care roles. But these come with severe ethical considerations. This behooves the question of the justification of these systems. As we argued before, the opacity-transparency relationship is normative and enables us to distinguish between justified and unjustified use cases.

AC was conceptualized, in the words of Rosalind Picard, to make human-computer interaction smooth and friction-free and nudge them towards better experiences [36]. This is especially when machines have become increasingly complex to allow easy maneuvering and understanding. Here we notice a version of the opacity-transparency relationship emerging at the conceptualization stage of AC. As machines become more complex and less procedurally transparent, they need access to user experience to function better for them. Put in terms of the opacity-transparency relationship: greater opacity of machines requires greater transparency of the humans in order to serve humans better. The actual shift in opacity-transparency though depends on the use case, design, and values embedded in it; greater human transparency may or may not always follow a concomitant reduction in the transparency of the machine. For example, an AC-based wearable that captures our heart rate is an extension of a medical-grade ECG and is procedurally transparent. But when integrated with Big Data or machine learning systems, it shifts towards greater opacity of machines and lesser opacity of human beings. Similarly, facial detection for driver drowsiness enables detection for the purposes of safety but still may constitute a high risk on the level of trustworthiness according to the new European AI Act that classifies AI applications based on the risk of their intended use [150].

Use cases may affect opacity and transparency at varied levels and to varied degrees. Their justification may then lie on whether its deployment is moral and protects fundamental values of justice, autonomy, and democracy. Transparency, as we have argued before, may be crucial to attaining certain instrumental and constitutive goods for human flourishing and welfare. The deployment of AC in healthcare settings like seizure detection and in communicating affects for individuals with autism spectrum disorders does precisely that. These affective systems, when localized, demonstrate an adequate concern for human opacity and enable human welfare. In other words, these systems balance opacity and transparency. To demonstrate the fuzzy nature of opacity in the opacity-transparency relationship, consider MACH. When it helps a candidate improve their interview skills by providing constructive feedback and a platform to retry and learn, it enhances their autonomy at the cost of their opacity. However, the same MACH software when used by an employer to gauge the mental state of a candidate before and during an interview when making a decision about their appointment violates their opacity and also discriminates on a protected feature. Similarly, the deployment of AC for workplace surveillance is unjustified as it reduces workers' opacity without any concomitant benefit to their welfare. Rather they hold possibilities of misrecognition that compromise the testimonial credibility of the workers. The deployment of facial recognition technology in security for gauging antisocial behavior or for purposes of policing to predict future crimes too affects opacity and raises severe concerns of privacy, compromised autonomy and requires careful balancing against the benefits of security. There are crucial concerns regarding the effectiveness of these systems [151] in predicting crime, whereas the implications of it on minorities and democracies are well captured [138]. The fact that the deployment of these technologies also goes with reduced accountability, they tend to tilt the opacity-transparency balance towards reduced opacity of humans with greater opacity of machines. It is thus important that affect recognition technologies be judged on their epistemic claims and the distributive consequences for which we have argued.

AC systems thus require proper calibration to find an adequate balance between the benefits of human transparency and opacity because reduced opacity comes at a significant cost to the user in terms of reduced autonomy, distributive burdens, manipulation, deception, and reduced credibility and, in some instances, have grave consequences for democracy. We thus argue that the shift towards greater transparency of the human ought to follow greater transparency of the machine and that transparency of the human should be of a kind that enhances individual autonomy and upholds democratic values of justice and equality.

B. Three Possible Objections

Before we conclude, we anticipate three possible objections to our argument. The first is that AC reduces human frustration and thus is convenient for human beings, and convenience has a moral weight that justifies the deployment of affective systems. Second, AC does not involve a loss of opacity, but privacy which depends on users' consent and the benefits they attribute to it. Third, our assessment is not sensitive to the actual powers of AC systems as they exist, but it tends to overplay them to derive conclusions that may not realize, at least not yet. Let us go through them one by one.

1) Convenience and Affective Computing: The promise of affective systems lies in reducing human frustration and nudging our responses in the desired direction. At the core of affective computing, thus lies the value of convenience and benefit. Let us call this the argument from convenience. It can thus be argued that convenience and efficiency have moral worth, and thus AC systems promoting them should thus be promoted. The argument could be framed in the following form: designers have moral reasons to make human interactions with machines as convenient and efficient as possible[2].

At the outset, the argument is both obvious and unproblematic. Individuals have both pragmatic and moral reasons to be less exposed to burdens and be offered options that enable overall happiness. This includes reduced frustration and better satisfaction. But beneficence and convenience arguments do not have an all-considered view; rather they treat human beings as utility maximizing agents. This is the thrust of the hypothetical example of a computer piano teacher that Rosalind Picard uses in one of her papers [152]. The computer piano teacher, like the best human teachers, through affect recognition, ensures skillful redirection and constitutes the difference between quitting or going on, akin to the teaching system that "tries to maximize pleasure and interest, while minimizing distress". In this example, when the computer nudges individuals towards learning by reducing frustration, it enables human functioning. But let's say if the same system is used to manipulate us into subscribing to a product or eliciting a 'behavioral trust' to ensure that the user returns to the technology time and again [153] it becomes a site of manipulation. This is the case with Facebook and other online platforms that manipulate individuals to optimize productivity and achieve efficient outcomes in pursuit of profits. Convenience here then comes at a cost to the user [154], i.e. to be able to experience the technology seamlessly, the user is required to hand over more data. This may come at the cost of diverse human interests [155] such as privacy [156], autonomy [157], transparency, accountability, justice, and equality. Similarly, systems like MACH can be justified as a pedagogical device that enables better preparation akin to an interview coach. But its use in actual interview scenarios structures interactions based on a certain predefined set and is used in a distributive scheme (like access to jobs) that raises serious concerns about distributive justice. Arguments of convenience thus have an 'all-considered view' and have a moral weight only in a society that already has the basic conditions of justice, transparency, and accountability [158].

The preference satisfaction view of AC that companies allude to while claiming to know individual wants better than the individuals themselves has a paternalistic dimension that treats affect detection as a commodity and human beings as utilitarian desire satisfaction machines who only care about the mental state of satisfaction. But human beings are not merely utility-maximizing agents who seek a mental state of less frustration and higher satisfaction. As Robert Nozick argues, individuals do not merely care about the experience of doing something but also want to do that thing, believing that some of them will turn out to be happy [159]. In other words, human beings enjoy the process (including going through failure and frustration) of performing tasks rather than merely seeking better experiences. To go back to the computer piano teacher example. It aids us till it does not compromise the processes through which human beings learn. The moment it starts offering only better outcomes, it undermines our capacity to exercise our freedom. Exercising autonomy requires that our choices are self-determined [160] and significant [161], that there are no undesirable interferences [162], and threats of domination and manipulation [163]. Human beings must be able to endorse and influence their own actions and decisions, deliberate over their options, and consider their background and context before finally making a choice [104]. In this regard, affective systems reducing human opacity are not necessarily justified even if it results in greater convenience for the human being. Considerations of convenience do tilt the balance towards reduced opacity of humans but are then undermined by weightier considerations of autonomy.

2) AC and Privacy: It can be argued that what AC impacts is a right to privacy, and by highlighting the loss as opacity, we misrecognize the nature and the extent of the problem. Privacy defendants may argue that emotion is a kind of information that the individual can choose or not to withhold based on the significance that they attach to it. Indeed concerns of privacy are intrinsic to the design of AC systems as they are invasive [164].

[2]This we notice in Picard's assertion we cited earlier. Albeit, she is careful in her assertion regarding the possible uses of this technology and the areas in which it should not tread and calls for a balanced approach between competing human interests [7]. But the fact that these assertions are already being made by individuals in the industry requires that they be clearly examined for their implications.

Behavior patterns of a person, their preferences, and their emotions about what matters to them or what frustrates them are all highly personal in nature [56]. It could be argued that issues of privacy do not account for the actual behavior of individuals who willingly submit themselves to technology. That technology should be judged on practices of consent that enable individuals to actually decide whether they wish to benefit from the service or not. Thus, AC interactions should rather be regulated through individual assessments of the costs and benefits and the weights that individuals attach to opacity and transparency.

This argument, though compelling, rests on the premise of equal background conditions for the exercise of consent. Practices of consent are based on asymmetric power relations [165] that frame the choices of the user. The user and the corporations offering these services are not equal. The costs of opting out may indeed be high for individuals when these services become essential. They may frame our sociality as Facebook does or are tied to accessing other services, like accessing job opportunities or refusing work opportunities (in cases of workplace surveillance). A refusal or failure to consent comes at a huge cost to the individual that they may not be willing to bear. The tradeoff thus operates under conditions of unequal power that frame the responses that individuals can offer. These tradeoffs also operate in an epistemically asymmetric environment whereby the function of the machine and the extent of our autonomy that is compromised are not always known to us. Individuals may willingly submit themselves to these systems thinking that they are credible. This we noticed in the case of the deployment of AC for the purpose of interviews that we have already highlighted in Section III-A and III-C2. Candidates are hardly aware of the functioning and effectiveness of these devices. Yet they ascribe greater credibility to it because the company tends to do so, and they feel that they are being judged by a superior entity. This is an epistemically asymmetrical position where individuals do not have enough information to engage in meaningful choices and are in no position to decide on the tradeoff.

A far more crucial point is that reading the loss as a mere loss of privacy would be missing the point. One may have reasons to allow access to emotions that may comply with legal stipulations of privacy despite the asymmetric conditions of power that underlie it. In this regard, privacy will be protected, and only affective systems that share affective information without the consent of the individual would violate the right to privacy. However, our claim is that affective systems undermine more than privacy. Let us understand this through a distinction between opacity and privacy.[3]

In literature, privacy is defined in its relation to other entities (space, time, and access). Koops et al. give a typology of privacy in which they use space and access as dimensions to understand privacy where they are situated on a continuum between restricted access to control [18]. This spectrum of privacy protects the rights of a person to remove access to their private spaces, such as their body, home, and even personal thoughts. It allows control over one's own data. Privacy thus has an informational and spatial character in terms of borders between the individual and the world [166] that includes mental information [167] that protects our inner world [27], [168]. Privacy and the ability to keep information secret are required for the well-being of the individual [169], their ability to self-determination [170], [171], and their capacity to choose [172].

Opacity, however, is an intrinsic unknowability of an entity. When applied to humans, it refers to the quality of keeping mental information such as thoughts, emotions, and opinions or physical information hidden unless they wish to make them explicit. There are aspects of humans that are unknowable because of their ontological constitution that makes machines incapable of knowing them.

Although machines today are able to capture a lot of information about human beings, it can be argued that much of this information is given out willingly or can be easily interpreted from the provided information. However, mental information which may be interpreted from affects crosses the boundary of our opacity. For example, affects such as sweating may be captured by wearables that can gauge galvanic skin conductance. This violates our ontological opacity with machines as they are able to gauge affects they could not before. However, when the wearable device concludes that we are sweating because we are nervous, this is an epistemic assumption that violates our mental opacity. At this point still, our privacy is maintained since this conclusion is localized. However, if this information is passed on to a party that may make important decisions based on it, such as an interviewer judging whether or not the person will get the job as they seem stressed or nervous in a critical setting, our privacy is now also violated. Similarly, consider an X-ray machine. This machine renders parts of the human body transparent which were previously opaque. This is a violation of opacity but not of privacy, as the patient consents to the use of the machine for diagnostic purposes. However, if this personal medical information is used for purposes the patient did not consent to, then it would be a violation of their privacy. AC enters into a space intrinsic to a human being and only knowable to an extent by another human being. This information is not private that allows discretion of whether or not to share it, but opaque because machines so far could not access it. This opacity was thus an intrinsic feature of human-machine interaction prior to AC because machines were not ontologically designed to access emotions. This relationship shifts with AC. In this regard, AC represents a loss of human opacity (in terms of access to emotions) but also undermines the control over affects and thus undermines secrecy (in terms of hiding the content of the affect or the emotion), a deliberate attempt to conceal it (privacy), and also misrepresent or obscure it from the gaze of the others.

3) The Capacity of Existing AC is Restrictive: A final consideration is due. It can be argued that our claim on human opacity and transparency by AC rests on a very expansive reading of emotion recognition as that of reading the innermost feelings of an individual [117]. What AC is capable of doing thus far is, in reality, quite restrictive. However, our argument still applies even with a more restrictive view of emotion recognition. Even if the goal of AC is restrictive, it still affects the opacity-transparency relationship and does raise concerns about privacy, manipulation, and deception. Similarly, epistemic arguments related to credibility too follow, not as a result of the actual design of AC, but rather from its uptake in the social spaces and mechanisms

[3]We thank the anonymous reviewer for asking us to clarify the distinction.

of distribution of credibility. The uptake of the technology in real-world settings is different from its actual capacities in a laboratory setting. AC with limited functioning can still be presented as being foolproof, having a higher epistemic status, and being more capable than it actually is. As Hupont et al. find in their study of facial processing applications 'that there are many "high-risk" applications in the market, even though some challenges are still to be solved to ensure that these systems are developed and evaluated in a trustworthy way according to the use case they will be used for, as required by the AI Act" [150] Governments, corporations, and social groups may have their own interests in deploying AC which may result in a social uptake where it can achieve a mythical status of having powers it does not have. This has unique implications for the users or those who are affected by such technology, and any ethical assessment of AC should not be devoid of an assessment of the circumstances and context in which it is deployed. This is because the ethical implications are not merely inherent in the design, but also in the way it interacts with its environment and human beings. When introduced in an inegalitarian institutional structure, it is bound to exacerbate existing social divisions and work to the benefit of those in power in the process compromising justice, human dignity, and democracy.

V. Conclusion

In this article, we have explored the ethical implications of AC systems through an opacity-transparency relationship by focusing on their epistemic assumptions and their impact on autonomy. We argued that AC systems shift the relationship towards greater transparency of human beings. This does not always follow a concomitant rise in machine transparency. Despite some obvious benefits in some use cases, AC systems have unique implications for human beings in terms of exposing them to deception, manipulation, and reduced autonomy, and raise crucial questions of justice. Affective technologies thus call for a need to balance their promises against the possibilities of abuse. Given unjust background conditions, this balancing cannot operate at an individual level but requires cooperative strategies that demand more transparency and accountability by ensuring liabilities from corporations, governments, and designers and better legal and regulatory structures, including laws that protect affective data. Current AC systems should thus be regulated by evaluating them on the opacity-transparency spectrum. While future affect recognition requires incorporating full disclosure of opacity-transparency at the design stage and need to be audited for their impact to ensure accountability. This requires collaborative and integrative effort between designers, ethicists, legal scholars, and civil society activists, and a vision for technology development that is not merely based on profits and efficiency but also on social good.

References

[1] M.-Z. Poh, K. Kim, A. Goessling, N. Swenson, and R. Picard, "Cardiovascular monitoring using earphones and a mobile device," *IEEE Pervasive Comput.*, vol. 11, no. 4, pp. 18–26, Fourth Quarter 2012.

[2] R. E. Kaliouby, R. Picard, and S. Baron-Cohen, "Affective computing and autism," *Ann. New York Acad. Sci.*, vol. 1093, no. 1, pp. 228–248, 2006.

[3] R. W. Picard and J. Klein, "Computers that recognise and respond to user emotion: Theoretical and practical implications," *Interacting Comput.*, vol. 14, no. 2, pp. 141–169, 2002.

[4] D. S. Messinger et al., "Affective computing, emotional development, and autism," in *The Oxford Handbook Affect. Comput.*. Oxford University Press, 2015. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780199942237.013.012

[5] Y. Koumpouros and T. Kafazis, "Wearables and mobile technologies in autism spectrum disorder interventions: A systematic literature review," *Res. Autism Spectr. Disord.*, vol. 66, 2019, Art. no. 101405.

[6] M. Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, and R. W. Picard, "MACH: My automated conversation coach," in *Proc. ACM Int. Joint Conf. Pervasive ubiquitous Comput.*, 2013, pp. 697–706.

[7] R. W. Picard, "Affective computing: Challenges," *Int. J. Hum.-Comput. Stud.*, vol. 59, no. 1/2, pp. 55–64, 2003.

[8] R. Cowie, "Ethical issues in affective computing," in *The Oxford Handbook of Affective Computing*. London, U.K.: Oxford Univ. Press, 2015, pp. 334–348.

[9] C. Reynolds and R. Picard, "Evaluation of affective computing systems from a dimensional metaethical position," in *Proc. First Augmented Cogn. Int. Conf.*, Las Vegas, NV., 2005.

[10] R. Picard, *Affective Computing Cambridge*. Cambridge, MA, USA: MIT Press, 1997.

[11] S. B. Daily et al., "Affective computing: Historical foundations, current applications, and future trends," in *Emotions and Affect in Human Factors and Human-Computer Interaction*. Amsterdam, The Netherlands: Elsevier, 2017, pp. 213–231.

[12] C. Reynolds and R. Picard, "Affective sensors, privacy, and ethical contracts," in *Proc. Extended Abstr. Hum. Factors Comput. Syst.*, 2004, pp. 1103–1106.

[13] J. Pitt, "Design contractualism for pervasive/affective computing," *IEEE Technol. Soc. Mag.*, vol. 31, no. 4, pp. 22–29, Winter 2012.

[14] N. Buitron and H. Steinmüller, "Governing opacity: Regimes of intention management and tools of legibility," *Ethnos*, vol. 2021, pp. 1–25, 2021. [Online]. Available: https://doi.org/10.1080/00141844.2021.2007154

[15] L. D. Introna, "Disclosive ethics and information technology: Disclosing facial recognition systems," *Ethics Inf. Technol.*, vol. 7, no. 2, pp. 75–86, Jun 2005. [Online]. Available: https://doi.org/10.1007/s10676--005-4583-2

[16] M. Heidegger, *Being and Time*. Oxford, U.K.: Blackwell, 1967. [Online]. Available: https://books.google.co.in/books?id=S57m5gW0L-MC

[17] M. Wheeler, "The reappearing tool: Transparency, smart technology, and the extended mind," *AI Soc.*, vol. 34, no. 4, pp. 857–866, 2019.

[18] B.-J. Koops, B. C. Newell, T. Timan, I. Skorvanek, T. Chokrevski, and M. Galic, "A typology of privacy," *Univ. Pennsylvania J. Int. Law*, vol. 38, 2016, Art. no. 483.

[19] Y. Van Den Eede, "In between us: On the transparency and opacity of technological mediation," *Found. Sci.*, vol. 16, no. 2, pp. 139–159, 2011.

[20] S. Turkle, *The Second Self: Computers and the Human Spirit*. Cambridge, MA, USA: MIT Press, 2005.

[21] É. Glissant, *Poetics of Relation*. Ann Arbor, MI, USA: University of Michigan Press, 1997.

[22] C. Birchall, "Introduction to 'secrecy and transparency' the politics of opacity and openness," *Theory, Culture Soc.*, vol. 28, no. 7/8, pp. 7–25, 2011.

[23] C. Birchall, *Radical Secrecy: The Ends of Transparency in Datafied America*. Minneapolis, MN, USA: Univ. Minnesota Press, 2021, vol. 60.

[24] P. Brey, "Ethical aspects of facial recognition systems in public places," *J. Inf. Commun. Ethics Soc.*, vol. 2, pp. 97–109, 2004.

[25] S. Bok, *Secrecy: On the Ethics of Concealment and Revelation*. New York, NY, USA: Oxford University Press, 1982.

[26] H. Surden and M.-A. Williams, "Technological opacity, predictability, and self-driving cars'(2016)," *Cardozo Law Rev.*, vol. 38, 2016, Art. no. 181.

[27] W. Hartzog and E. Selinger, "Surveillance as loss of obscurity," *Washington Lee Law Rev.*, vol. 72, 2015, Art. no. 1343.

[28] T. Nagel, "Concealment and exposure," *Philosophy Public Affairs*, vol. 27, no. 1, pp. 3–30, 1998.

[29] O. O'neill, *A Question of Trust: The BBC Reith Lectures 2002*. Cambridge, U.K.: Cambridge Univ. Press, 2002.

[30] Z. Blas, "Informatic opacity," *Posthuman Glossary*, R. Braidotti and M. Hlavajova, Eds. Bloomsbury Academic, 2018, pp. 103–114.

[31] J. Waldron, "It's all for your own good," 2014. [Online]. Available: https://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/

[32] D. Santoro and M. Kumar, *Speaking Truth to Power-A Theory of Whistleblowing*, vol. 6. Berlin, Germany: Springer, 2018.

[33] R. Binns, "Algorithmic accountability and public reason," *Philosophy Technol.*, vol. 31, no. 4, pp. 543–556, 2018.

[34] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. Zhu, "Explainable AI: A brief survey on history, research areas, approaches and challenges," in *Proc. Int. Conf. Natural Lang. Process. Chin. Comput.*, Springer, 2019, pp. 563–574.

[35] P. Brey, "Disclosive computer ethics," *ACM Sigcas Comput. Soc.*, vol. 30, no. 4, 2000, pp. 10–16.

[36] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 2000.

[37] S. D'Mello, "Affective/emotional computing," in *Encyclopedia of Sciences and Religions*. Berlin, Germany: Springer, 2013, pp. 29–31.

[38] R. Calvo, S. D'Mello, J. Gratch, and A. Kappas, "The oxford handbook of affective computing," in *Oxford Library of Psychology*. London, U.K.: Oxford Univ. Press, 2015.

[39] A. R. Damasio, "Descartes' error and the future of human life," *Sci. Amer.*, vol. 271, no. 4, pp. 144–144, 1994.

[40] A. M. Isen, "Positive affect and decision making," *Handbook of Emotions 2nd Ed.*, M. Lewis and J. M. Haviland, Eds. Guilford Press, 2000, pp. 417–435.

[41] J. Klein, Y. Moon, and R. W. Picard, "This computer responds to user frustration: Theory, design, and results," *Interacting Comput.*, vol. 14, no. 2, pp. 119–140, 2002.

[42] J. Scheirer, R. Fernandez, J. Klein, and R. W. Picard, "Frustrating the user on purpose: A step toward building an affective computer," *Interacting Comput.*, vol. 14, no. 2, pp. 93–118, 2002.

[43] R. Heersmink, "Extended mind and cognitive enhancement: Moral aspects of cognitive artifacts," *Phenomenol. Cogn. Sci.*, vol. 16, no. 1, pp. 17–32, 2017.

[44] A. Clark and D. Chalmers, "The extended mind," *Analysis*, vol. 58, no. 1, pp. 7–19, 1998.

[45] K. Boehner, R. DePaula, P. Dourish, and P. Sengers, "How emotion is made and measured," *Int. J. Hum.-Comput. Stud.*, vol. 65, no. 4, pp. 275–291, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581906001844

[46] A. Ben-Ze'ev, *The Subtlety of Emotions*. Cambridge, MA, USA: MIT Press, 2001.

[47] R. Berrios, P. Totterdell, and S. Kellett, "Silver linings in the face of temptations: How mixed emotions promote self-control efforts in response to goal conflict," *Motivation Emotion*, vol. 42, no. 6, pp. 909–919, 2018.

[48] A. Scarantino and R. De Sousa, "Emotion," in *The Stanford Encyclopedia of Philosophy*, Summer 2021 ed., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2021.

[49] P. Mantello, M.-T. Ho, M.-H. Nguyen, and Q.-H. Vuong, "Bosses without a heart: Socio-demographic and cross-cultural determinants of attitude toward emotional ai in the workplace," *AI Soc.*, vol. 38, pp. 97–119, 2023.

[50] B. Parkinson, "Do facial movements express emotions or communicate motives?," *Pers. Social Psychol. Rev.*, vol. 9, no. 4, pp. 278–311, 2005.

[51] C. Frith, "Role of facial expressions in social interactions," *Philos. Trans. Roy. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3453–3458, 2009.

[52] R. Cowie, "Perceiving emotion: Towards a realistic understanding of the task," *Philos. Trans. Roy. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3515–3525, 2009.

[53] IEEE, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Ed., 2019. [Online]. Available: https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

[54] B. De Gelder, "Why bodies? Twelve reasons for including bodily expressions in affective neuroscience," *Philos. Trans. Roy. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3475–3484, 2009.

[55] R. Cowie, "What are people doing when they assign everyday emotion terms?," *Psychol. Inquiry*, vol. 16, no. 1, pp. 11–15, 2005.

[56] S. Richardson, "Affective computing in the modern workplace," *Bus. Inf. Rev.*, vol. 37, no. 2, pp. 78–85, 2020.

[57] D. C. Richardson, "Frown interpretation and perception: An investigation of the appearance of negative affect," *Perception*, vol. 48, no. 6, pp. 505–535, 2019.

[58] K. Crawford et al., "AI now 2019 report," 2019. [Online]. Available: https://ainowinstitute.org/publication/ai-now-2019-report-2

[59] K. Crawford et al., "Time to regulate AI that interprets human emotions," *Nature*, vol. 592, no. 7853, pp. 167–167, 2021.

[60] A. X. Wu, "The ambient politics of affective computing," *Public Culture*, vol. 34, no. 1, pp. 21–45, 2022.

[61] L. Stark and J. Hoey, "The ethics of emotion in artificial intelligence systems," in *Proc. ACM Conf. Fairness Accountability Transparency*, ser. FAccT '21. New York, NY, USA, 2021, pp. 782–793. [Online]. Available: https://doi.org/10.1145/3442188.3445939

[62] L. F. Barrett, *How Emotions are Made: The Secret Life of the Brain*. Boston, MA, USA: Houghton Mifflin, 2017.

[63] P. Ekman, "Basic emotions," *Handbook Cogn. Emotion*, vol. 98, no. 45-60, 1999, Art. no. 16.

[64] G. W. Boyd, "The body, its emotions, the self, and consciousness," *Perspectives Biol. Med.*, vol. 55, no. 3, pp. 362–377, 2012.

[65] Q.-H. Vuong, "Emotions as dynamic processes: A review of theory and research," *Int. J. Res. Bus. Social Sci.*, vol. 8, no. 2, pp. 23–35, 2019.

[66] A. McStay, *Emotional AI: The Rise of Empathic Media*. Newbury Park, CA, USA: Sage, 2018.

[67] A. F. Beavers and J. P. Slattery, "On the moral implications and restrictions surrounding affective computing," in *Emotions and Affect in Human Factors and Human-Computer Interaction*. Amsterdam, The Netherlands: Elsevier, 2017, pp. 143–161.

[68] J. Bullington, "'Affective'computing and emotion recognition systems: The future of biometric surveillance?," in *Proc. 2nd Annu. Conf. Inf. Secur. Curriculum Develop.*, 2005, pp. 95–99.

[69] A. Carnevale, "Will robots know us better than we know ourselves?," *Robot. Auton. Syst.*, vol. 86, pp. 144–151, 2016.

[70] Y. N. Harari and T. Harris, "When tech knows you better than you know yourself," *Wired*, Oct. 2018. [Online]. Available: https://www.wired.com/story/artificial-intelligenceyuval-noah-harari-tristan-harris

[71] J. Jacobs, "Neural networks allow us to 'read faces' in a new way," 2018. [Online]. Available: https://www.ft.com/content/b387b1a2-de7b-11e7-a0d4--0944c5f49e46

[72] D. Sperber and D. Wilson, "Remarks on relevance theory and the social sciences," *Multilingua*, vol. 16, pp. 145–152, Jan. 1997.

[73] M. Gonzalez-Franco, G. D. Clemenson, and A. Miller, "How GPS weakens memory—and what we can do about it," 2021. [Online]. Available: https://www.scientificamerican.com/article/how-gps-weakens-memory-mdash-and-what-we-can-do-about-it/

[74] D. J. Pauleen, R. Evaristo, R. M. Davison, S. Ang, M. Alanis, and S. Klein, "Cultural bias in information systems research and practice: Are you coming from the same place I am?," *Commun. Assoc. Inf. Syst.*, vol. 17, no. 1, 2006, Art. no. 17.

[75] P. Ratner, "How to beat A.I. in landing a job," 2020. [Online]. Available: https://bigthink.com/the-future/how-to-beat-a-i-in-landing-a-job/rebelltitem2

[76] Z. Jaser and D. Petrakaki, "Are you prepared to be interviewed by an AI?," *Harvard Bus. Rev.*, 2023. [Online]. Available: https://hbr.org/2023/02/are-you-preparedto-be-interviewed-by-an-ai

[77] Z. Jaser, D. Petrakaki, R. Starr, and E. Oyarbide-Magana, "Where automated job interviews fall short," *Harvard Bus. Rev.*, vol. 100, no. 1, pp. 122–131, 2022.

[78] F. Hirt, E. Werlen, I. Moser, and P. Bergamin, "Measuring emotions during learning: Lack of coherence between automated facial emotion recognition and emotional experience," *Open Comput. Sci.*, vol. 9, no. 1, pp. 308–317, 2019.

[79] K. Ball, "Workplace surveillance: An overview," *Labor Hist.*, vol. 51, pp. 87–106, 2010.

[80] K. Ball, "Exposure," *Inf., Commun. Soc.*, vol. 12, no. 5, pp. 639–657, 2009. [Online]. Available: https://doi.org/10.1080/13691180802270386

[81] M. Fricker, *Epistemic Injustice: Power and the Ethics of Knowing*. London, U.K.: Oxford Univ. Press, 2007.

[82] R. Kamtekar, "Situationism and virtue ethics on the content of our character," *Ethics*, vol. 114, no. 3, pp. 458–491, 2004.

[83] E. Goffman, *The Presentation of Self in Everyday Life*. New York, NY, USA: Doubleday, 1978, pp. 1–10.

[84] H. Arendt, "What was authority?," *NOMOS: Amer. Soc. Political Legal Philosophy*, vol. 1, 1958, Art. no. 81.

[85] S. Delacroix and M. Veale, "Smart technologies and our sense of self: Going beyond epistemic counter-profiling," in *Life and the Law in the Era of Data-Driven Agency*. Cheltenham, U.K.: Edward Elgar Publishing, 2020.

[86] A. J. Julius, "Basic structure and the value of equality," *Philosophy Public Affairs*, vol. 31, no. 4, pp. 321–355, 2003.

[87] S. Barsade and O. A. O'Neill, "Manage your emotional culture," *Harvard Bus. Rev.*, vol. 94, no. 1, pp. 58–66, 2016.

[88] M. Scheutz, "13 the inherent dangers of unidirectional emotional bonds between humans and social robots," in *Robot Ethics: The Ethical and Social implications of robotics*. Cambridge, MA, USA: MIT Press, 2011, Art. no. 205.

[89] M. Scheutz, "The affect dilemma for artificial agents: Should we develop affective artificial agents?," *IEEE Trans. Affect. Comput.*, vol. 3, no. 4, pp. 424–433, Fourth Quarter 2012.

[90] R. Cowie, "The good our field can hope to do, the harm it should avoid," *IEEE Trans. Affect. Comput.*, vol. 3, no. 4, pp. 410–423, Fourth Quarter 2012.

[91] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: A survey," *Int. J. Social Robot.*, vol. 5, pp. 291–308, 2013.

[92] C.-H. Wu, Y.-M. Huang, and J.-P. Hwang, "Review of affective computing in education/learning: Trends and challenges," *Brit. J. Educ. Technol.*, vol. 47, no. 6, pp. 1304–1323, 2016. [Online]. Available: https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12324

[93] T. Körtner, "Ethical challenges in the use of social service robots for elderly people," *Zeitschrift für Gerontologie und Geriatrie*, vol. 49, no. 4, pp. 303–307, 2016.

[94] A. Pirni, M. Balistreri, M. Capasso, S. Umbrello, and F. Merenda, "Robot care ethics between autonomy and vulnerability: Coupling principles and practices in autonomous systems for care," *Front. Robot. AI*, vol. 8, 2021, Art. no. 184.

[95] V. Held et al., *Feminist Morality: Transforming Culture, Society, and Politics*. Chicago, IL, USA: Univ. Chicago Press, 1993.

[96] V. Held et al., *The Ethics of Care: Personal, Political, and Global*. London, U.K.: Oxford Univ. Press, 2006.

[97] N. Noddings, *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley, CA, USA: Univ. California Press, 1984.

[98] C. Gilligan, *In A Different Voice: Psychological Theory and Women's Development*, vol. 326. Cambridge, MA, USA: Harvard University Press, 1982.

[99] J. C. Tronto, *Moral Boundaries: A Political Argument for an Ethic of Care*. London, U.K.: Psychology Press, 1993.

[100] I. Wachsmuth, "Robots like me: Challenges and ethical issues in aged care," *Front. Psychol.*, vol. 9, 2018, Art. no. 432.

[101] R. Noggle, "The Ethics of Manipulation," in *The Stanford Encyclopedia of Philosophy*, Summer 2 ed., E. N. Zalta, Ed., Metaphysics Research Lab, Stanford, CA, USA: Stanford University, 2022.

[102] R. E. Goodin, "Manipulatory politics," Yale University Press, 1980.

[103] T. Scanlon, "What we owe to each other," Harvard University Press, 1998. [Online]. Available:http://www.jstor.org/stable/j.ctv134vmrn

[104] D. Susser, "Invisible influence: Artificial intelligence and the ethics of adaptive choice architectures," in *Proc. AAAI/ACM Conf. AI Ethics Soc.*, 2019, pp. 403–408.

[105] A. Veltman and M. Piper, *Autonomy, Oppression, and Gender*. London, U.K.: Oxford Univ. Press, 2014.

[106] R. Calo, "Digital market manipulation," vol. 82, no. 4, pp. 995–1040, 2014. [Online]. Available: https://ssrn.com/abstract=2309703

[107] S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019*. London, U.K.: Profile books, 2019.

[108] B. Frischmann and E. Selinger, *Re-Engineering Humanity*. Cambridge, U.K.: Cambridge Univ. Press, 2018.

[109] M. Kaptein and D. Eckles, "Selecting effective means to any end: Futures and ethics of persuasion profiling," in *Proc. Int. Conf. Persuasive Technol.*, Springer, 2010, pp. 82–93.

[110] A. D. Angie, S. Connelly, E. P. Waples, and V. Kligyte, "The influence of discrete emotions on judgement and decision-making: A meta-analytic review," *Cogn. Emotion*, vol. 25, no. 8, pp. 1393–1422, 2011.

[111] S. Steinert and O. Friedrich, "Wired emotions: Ethical issues of affective brain–computer interfaces," *Sci. Eng. Ethics*, vol. 26, no. 1, pp. 351–367, 2020.

[112] C. Arthur, "Facebook emotion study breached ethical guidelines, researchers say," *The Guardian*, vol. 30, 2014.

[113] L. Munn, "Angry by design: Toxic communication and technical architectures," *Humanities Social Sci. Commun.*, vol. 7, no. 1, pp. 1–11, 2020.

[114] A. McStay, "Empathic media and advertising: Industry, policy, legal and citizen perspectives (the case for intimacy)," *Big Data Soc.*, vol. 3, no. 2, 2016, Art. no. 2053951716666868.

[115] S. Zhao, S. Wang, M. Soleymani, D. Joshi, and Q. Ji, "Affective computing for large-scale heterogeneous multimedia data: A survey," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 15, no. 3s, pp. 1–32, 2019.

[116] M. He and A. Li, "The impact of sad-spending on emotional recovery process," *Psychology*, vol. 7, no. 1, pp. 85–91, 2016.

[117] G. Greene, "The Ethics of AI and Emotional Intelligence," Tech. Rep., 2020. [Online]. Available: https://partnershiponai.org/wp-content/uploads/2021/08/PAI_The-ethics-of-AI-and-emotional-intelligence_073020.pdf

[118] B. Horowitz, "Mobile health apps leak sensitive data through APIs, report finds," 2021. [Online]. Available: https://www.fiercehealthcare.com/tech/mobile-health-apps-leak-sensitive-data-through-apis-report-finds

[119] S. Pelley, "Whistleblower: Facebook is misleading the public on progress against hate speech, violence, misinformation," 2021. [Online]. Available: https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021--10-03/

[120] S. Vaidhyanathan, *Antisocial Media: How Facebook Disconnects us and Undermines Democracy*. London, U.K.: Oxford Univ. Press, 2018.

[121] F. Zuiderveen Borgesius et al., "Online political microtargeting: Promises and threats for democracy," *Utrecht Law Rev.*, vol. 14, no. 1, pp. 82–96, 2018. [Online]. Available: https://ssrn.com/abstract=3128787

[122] K. Yeung, "'hypernudge': Big data as a mode of regulation by design," *Inf., Commun. Soc.*, vol. 20, no. 1, pp. 118–136, 2017.

[123] J. H. Søraker and P. Brey, "Ambient intelligence and problems with inferring desires from behaviour," *Int. Rev. Inf. Ethics*, vol. 8, pp. 7–12, 2007.

[124] Z. Drayson and A. Clark, "Cognitive disability and embodied, extended minds," Adam Cureton, and David T. Wasserman (eds), *The Oxford Handbook of Philosophy and Disabil.*, Oxford Handbooks. Accessed: May 24, 2023. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780190622879.013.10

[125] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, "Affective computing and sentiment analysis," in *A Practical Guide to Sentiment Analysis*. Berlin, Germany: Springer, 2017, pp. 1–10.

[126] L. Urquhart, A. Laffer, and D. Miranda, "Working with affective computing: Exploring uk public perceptions of ai enabled workplace surveillance," 2022, *arXiv:2205.08264*.

[127] K. Ball, "Organization, surveillance and the body: Towards a politics of resistance," *Organization*, vol. 12, no. 1, pp. 89–108, 2005.

[128] G. La Torre, A. Esposito, I. Sciarra, and M. Chiappetta, "Definition, symptoms and risk of techno-stress: A systematic review," *Int. Arch. Occup. Environ. Health*, vol. 92, no. 1, pp. 13–35, 2019.

[129] G. Semi, "Alex rosenblat, uberland. how algorithms are rewriting the rules of work, oakland, university of california press, 2018, 272 pp," *Rassegna Ital. di Sociologia*, vol. 61, no. 3, pp. 671–674, 2020.

[130] D.-Y. Jeung, C. Kim, and S.-J. Chang, "Emotional labor and burnout: A review of the literature," *Yonsei Med. J.*, vol. 59, no. 2, pp. 187–193, 2018.

[131] Y. Gu and X. You, "Recovery experiences buffer against adverse well-being effects of workplace surface acting: A two-wave study of hospital nurses," *J. Adv. Nurs.*, vol. 76, no. 1, pp. 209–220, 2020.

[132] V. Gabrielle, "Gamified Life," 2018. [Online]. Available: https://aeon.co/essays/how-employers-have-gamified-work-for-maximum-profit

[133] A. Marciano, "Reframing biometric surveillance: From a means of inspection to a form of control," *Ethics Inf. Technol.*, vol. 21, no. 2, pp. 127–136, 2019.

[134] J. Rawls, "A theory of justice," in *A Theory of Justice*. Cambridge, MA, USA: Harvard Univ. Press, 1999.

[135] D. Belkin, "Colleges mine data on their applicants," 2019. [Online]. Available: https://www.wsj.com/articles/the-data-colleges-collect-on-applicants-11548507602

[136] S. Ji, S. Pan, X. Li, E. Cambria, G. Long, and Z. Huang, "Suicidal ideation detection: A review of machine learning methods and applications," *IEEE Trans. Comput. Social Syst.*, vol. 8, no. 1, pp. 214–226, Feb. 2021.

[137] F. Eyben et al., "Emotion on the road—necessity, acceptance, and feasibility of affective computing in the car," *Adv. Hum.-Comput. Interact.*, vol. 2010, pp. 1–17, 2010.

[138] E. Smith, E. A. Storch, H. Lavretsky, J. L. Cummings, and H. A. Eyre, "Affective computing for brain health disorders," in *Handbook of Computational Neurodegeneration*. Berlin, Germany: Springer, 2021, pp. 1–14.

[139] E. M. Kleiman et al., "Digital phenotyping of suicidal thoughts," *Depression Anxiety*, vol. 35, no. 7, pp. 601–608, 2018.

[140] A. Sano et al., "Identifying objective physiological markers and modifiable behaviors for self-reported stress and mental health status using wearable sensors and mobile phones: Observational study," *J. Med. Internet Res.*, vol. 20, no. 6, 2018, Art. no. e9410.

[141] K. Lippert-Rasmussen, ""we are all different": Statistical discrimination and the right to be treated as an individual," *J. Ethics*, vol. 15, no. 1, pp. 47–59, 2011.

[142] R. J. Arneson, "What is wrongful discrimination," *San Diego Law Rev.*, vol. 43, 2006, Art. no. 775.

[143] F. Schauer, *Profiles, Probabilities, and Stereotypes*. Cambridge, MA, USA: Harvard Univ. Press, 2006.

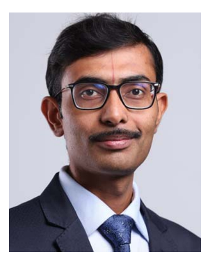
[144] S. Barocas and A. D. Selbst, "Big data's disparate impact," *California Law Rev.*, vol. 104, 2016, Art. no. 671.

[145] M. Leese, "The new profiling: Algorithms, black boxes, and the failure of anti-discriminatory safeguards in the european union," *Secur. Dialogue*, vol. 45, no. 5, pp. 494–511, 2014.
[146] M. Hildebrandt, *Smart Technologies and the End (s) of Law: Novel Entanglements of Law and Technology*. Cheltenham, U.K.: Edward Elgar Publishing, 2015.
[147] I. Manokha, "Facial analysis AI is being used in job interviews–it will probably reinforce inequality," *The Conversation*, vol. 7, 2019. [Online]. Available: https://theconversation.com/facial-analysis-ai-is-being-used-in-job-interviews-it-will-probably-reinforce-inequality-124790
[148] P. Bourdieu, "The forms of capital. cultural theory: An anthology," in *Handbook of Theory and Research for the Sociology of Education*. Westport, CT, USA: Greenwood Press, 1986, pp. 241–258.
[149] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data Soc.*, vol. 3, no. 1, 2016, Art. no. 2053951715622512.
[150] I. Hupont, S. Tolan, H. Gunes, and E. Gómez, "The landscape of facial processing applications in the context of the european AI act and the development of trustworthy systems," *Sci. Rep.*, vol. 12, no. 1, 2022, Art. no. 10688.
[151] V. L. Raposo, "When facial recognition does not 'recognise': Erroneous identifications and resulting liabilities," *AI & SOCIETY*, 2023. [Online]. Available: https://doi.org/10.1007/s00146--023-01634-z
[152] R. W. Picard, *Affective Computing*, vol. 2139. Cambridge, MA, USA: MIT Press, 1995, Art. no. 92.
[153] M. Beard, "Big tech's trojan horse to win your trust," 2020. [Online]. Available: https://ethics.org.au/big-techs-trojan-horse-to-win-your-trust/
[154] N. Kobie, "What is the Internet of Things," *The Guardian*, vol. 6, 2015.
[155] D. De Cremer and G. Kasparov, "AI should augment human intelligence, not replace it," *Harvard Bus. Rev.*, Mar. 2021. [Online]. Available: https://hbr.org/2021/03/ai-shouldaugment-human-intelligence-not-replace-it
[156] R. Wang, "Beware trading privacy for convenience," *Harvard Bus. Rev.*, Jun. 2013. [Online]. Available: https://hbr.org/2013/06/beware-trading-privacyfor-convenience
[157] E. J. Ottensmeyer and M. A. Heroux, "Ethics, public policy, and managing advanced technologies: The case of electronic surveillance," *J. Bus. Ethics*, vol. 10, no. 7, pp. 519–526, 1991.
[158] J. B. Wight, "The ethics behind efficiency," *J. Econ. Educ.*, vol. 48, no. 1, pp. 15–26, 2017.
[159] R. Nozick, *Anarchy, State, and Utopia*. New York, NY, USA: Basic Books, 1974.
[160] J. Raz, *The Morality of Freedom*. Oxford, U.K.: Clarendon Press, 1986.
[161] T. M. Scanlon, "Responsibility and the value of choice," *Think*, vol. 12, no. 33, pp. 9–16, 2013.
[162] I. Berlin, "Two concepts of liberty," in *Four Essays on Liberty*. London, U.K.: Oxford Univ. Press 1969.
[163] P. Pettit, "Freedom as antipower," *Ethics*, vol. 106, no. 3, pp. 576–604, 1996.
[164] N. M. Yusoff and S. S. Salim, "Shared mental model processing in visualization technologies: A review of fundamental concepts and a guide to future research in human-computer interaction," in *Proc. Int. Conf. Hum.-Comput. Interact.*, Springer, 2020, pp. 238–256.
[165] T. Monahan, "Visualizing the surveillance archive: Critical art and the dangers of transparency," Monahan, Torin. in press. Visualizing the Surveillance Archive: Critical Art and the Dangers of Transparency. In Law and the Visible, edited by A. Sarat, L. Douglas and MM Umphrey.: University of Massachusetts Press, pp. 130–155, 2020. [Online].Available: https://ssrn.com/abstract=3816933 or http://dx.doi.org/10.2139/ssrn.3816933
[166] G. T. Marx, "Murky conceptual waters: The public and the private," *Ethics Inf. Technol.*, vol. 3, no. 3, pp. 157–169, 2001.
[167] M. Ienca and G. Malgieri, "Mental data protection and the GDPR," *J. Law Biosci.*, vol. 9, no. 1, 2022, Art. no. lsac006.
[168] D. J. Solove, *The Digital Person: Technology and Privacy in the Information Age*, vol. 1. New York, NY, USA: New York Univ. Press, 2004.
[169] P. B. Newell, "A systems model of privacy," *J. Environ. Psychol.*, vol. 14, no. 1, pp. 65–78, 1994.
[170] P. K. Masur, "How online privacy literacy supports self-data protection and self-determination in the age of information," *Media Commun.*, vol. 8, no. 2, pp. 258–269, 2020.
[171] A. Rouvroy and Y. Poullet, "The right to informational self-determination and the value of self-development: Reassessing the importance of privacy for democracy," in *Reinventing Data Protection?* Berlin, Germany: Springer, 2009, pp. 45–76.
[172] J. van den Hoven, M. Blaauw, W. Pieters, and M. Warnier, "Privacy and Information Technology," in *The Stanford Encyclopedia of Philosophy*, Summer 2020 ed., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2020.

Manohar Kumar received the PhD degree in political theory from LUISS University, Rome, in 2013 and author of 'Speaking Truth to Power: A Theory of Whistleblowing'. He held postdoctoral fellowships with IIT Delhi and AMSE, Aix Marseille University, and is interested in Moral and Political Philosophy and issues related to whistleblowing, civil disobedience, democratic secrecy, epistemic injustice, surveillance, digital citizenship, and dissent.

Aisha Aijaz received the MS degree in data science from Heriot-Watt University. She is currently working toward the PhD degree with the Indraprastha Institute of Information Technology Delhi. Her research includes embedding ethics into artificial intelligence systems using ontology modeling and knowledge representation. She is currently working on building an ontology that may facilitate the moral reasoning of AI systems.

Omkar Chattar is currently working toward the PhD degree with Indraprastha Institute of Information Technology Delhi, interested in the ethics of technology and artificial intelligence, with special interests in existentialism, phenomenology, feminist ethics, queer feminism, postphenomenology, and ethics. Currently working on ethics of social media and use of chatbots as detterent for hate speech.



Jainendra Shukla (Member, IEEE) received the BE degree in IT from the University of Mumbai in First Class with Distinction in 2009, the MTech degree in IT with Specialization in Robotics from IIIT-Allahabad in 2012, and the PhD degree with Industrial Doctorate Distinction and International Doctorate Distinction from Universitat Rovira i Virgili, Spain, in 2018. His research interests involve empowering machines with emotional intelligence and adaptive interaction ability to improve the quality of life in health and social care. He co-developed the first-ever MOOC course on Affective Computing for NPTEL.

Raghava Mutharaju received the PhD degree in computer science and engineering from Wright State University, USA, in 2016. He is an assistant professor with the Computer Science and Engineering department of IIIT-Delhi, India and leads the Knowledgeable Computing and Reasoning (KRaCR; pronounced as cracker) Lab. He has worked in Industry research labs such as GE Research, IBM Research, Bell Labs, and Xerox Research. His research interest is in Semantic Web and, in general, in knowledge representation and reasoning. This includes knowledge graphs, ontology modelling, reasoning, querying, and its applications. He has published at several venues, such as ISWC, ESWC, ECAI, and WISE. He has co-organized workshops at ISWC 2020, WWW 2019, WebSci 2017, ISWC 2015 and tutorials at ISWC 2019, IJCAI 2016, AAAI 2015 and ISWC 2014. He has been on the Program Committee of several (Semantic) Web conferences such as AAAI, WWW, ISWC, ESWC, CIKM, K-CAP and SEMANTiCS. More information is available on his lab's homepage at https://kracr.iiitd.edu.in/.