ARTICLE

MIT Science Policy Review

Should machines be allowed to 'read our minds'? Uses and regulation of biometric techniques that attempt to infer mental states

Constanza M. Vidal Bustamante[1,*], Karolina Alama-Maruta[2], Carmen Ng[3], and Daniel D.L. Coppersmith[1]

Edited by Christopher Miller and Grant A. Knappe

HIGHLIGHTS

- Biometric data are increasingly used to attempt to infer individuals' momentary emotional and cognitive states, like stress and fatigue, as well as intentions, preferences, and health status
- Use cases range from personal wellness tracking and clinical monitoring in research settings to the surveillance of students and workers
- Many commercialized uses of biometrics for mental inference have limited scientific support and pose important ethical concerns related to individuals' privacy and self-determination
- Rigorous scientific research, more precise and proactive legal protections, and sustained global cooperation will be important to access the benefits of biometric techniques in fields like healthcare while mitigating individual and societal harms

Biometric data, such as facial expressions, voice, and heart rate, are increasingly used to make inferences about individuals' momentary emotional and cognitive states, like stress and fatigue, and for the categorization of more stable mental features, like intentions, preferences, and health status. This review provides an overview and discussion of common biometric techniques that attempt to infer such mental states, their technical and ethical challenges, and the current regulatory landscape. Drawing from use cases in personal fitness and wellness tracking, clinical research, and the monitoring of students and workers, we show that although these techniques promise greater objectivity, efficiency, and accuracy in the assessment of mental states and related decision-making, scientific evidence for these claims remains limited. The widespread use of these techniques, especially outside of regulated research settings, poses important technical and ethical challenges – from the exposure of highly sensitive biometric data to breaches and exploitation, to privacy violations and the use of faulty inferences to make consequential judgments about individuals' qualities. We review the strengths and limitations of current legislation in Europe and the United States related to biometric techniques, and present general considerations for regulation moving forward. We conclude that accessing the benefits of biometric techniques (e.g., for consensual clinical monitoring and care) while guarding against their harms may require rigorous scientific research, more precise and proactive legal protections, and sustained global cooperation.

Recent advances in sensing technologies, artificial intelligence (AI), and big data analytics have powered the development and spread of biometric techniques, making it easier to collect and process data related to physical, physiological, and behavioral features of the human body. While public discussion and regulation of biometric techniques tends to focus on the *identification* of individuals via first-generation biometrics like face geometry and fingerprints, considerably less attention has been paid to the use of biometrics for the inference of individuals' *mental states*.

Second-generation or *behavioral* biometrics, including voice, gait, facial expressions, heart rate, and brain activity, are increasingly leveraged in the attempt to detect individuals' momentary emotional and cognitive states, like stress and fatigue, and for the categorization of more stable mental features, like intentions, preferences, and health status [1]. In a striking example, customs and border control agencies around the world have used facial expression data to surmise intentions and identify 'dangerous' individuals. While these systems have been condemned due to human rights violations and their largely unscientific basis, the use of 'smart borders' was once under serious consideration in Europe, and it remains common practice in China's railway and subway stations and during police interrogations [2].

Biometric techniques claiming to detect mental states are in fact increasingly embedded in daily life (see Figure 1 for example use cases across several fields). In addition to watches, rings, and headbands estimating individuals' stress response, now cameras capturing facial expressions and sensor-bearing everyday objects measuring brain activity and posture are also becoming common. Critically, the intended

[1] Department of Psychology, Harvard University, Cambridge, MA
[2] University of Warsaw, Warsaw, Poland
[3] Technical University of Munich, Munich, Germany
* Email: cvidal@g.harvard.edu

The authors declare no conflict of interest.

Field	Purpose	Examples of sensing tools and data streams	Examples of specific inference
Personal trackers	Tracking personal trends in fitness, sleep, mood, wellness	Smartphone apps and wearable devices (wristbands, headbands, rings, etc.) capturing GPS location, activity, temperature, heart rate, electrodermal activity, voice, movement, gait	Stress response, mood
Health	Monitoring clinical symptoms and informing treatment response Continuous monitoring for at-risk populations, such as the elderly, patients with suicidal ideation, patients in intensive care units	Smartphones and wearables (see example data streams above) for analysis of locations visited, social behavior, mood, sleep, physical activity, etc. Ambient cameras	Clinical risk states (e.g., risk of schizophrenia relapse, suicidal ideation, cognitive impairment)
Education	Track student engagement in class Ranking of students' cognitive states and performance; assessing teachers' performance	Ambient cameras capturing facial expression and posture	Attentiveness
Work	Screening job candidates during interviews Tracking work performance and engagement	Polygraphs; cameras capturing facial expression and posture, microphones for voice (tone, prosody, etc.) and linguistic analysis Smart cushions, caps, other wearables capturing presence at desk, brain activity, heart rate	Job fit, skills, and projected performance Fatigue, attentiveness; anxiety, rage, sadness
Law enforcement	Detecting suspicious behavior and intentions to commit a crime, e.g., at border and migration control	Polygraphs; cameras for analysis of micro-expressions, such as eye blink, increase in face redness or head movement	Stress and anxiety potentially associated with lying
Transportation	Semi-autonomous vehicles: Assess whether driver / pilot is in appropriate state to drive / fly	Eye tracking and iris scans, smartphone sensors capturing GPS location and speed	Fatigue, attentiveness, anger, intoxication
Gaming	Control gaming activity via user's gaze or brain activity	Camera with eye-tracking capabilities and headset with EEG electrodes	Intended action (e.g., walk, run, jump, etc.)

Figure 1: Example uses of biometric technologies that attempt to infer mental states.

uses range from seemingly benign, such as personal wellness tracking and clinical symptom monitoring to study and treat mental illness, to more insidious and controversial, including real-time surveillance of students' and workers' engagement and performance. The scale, speed, and scope of biometric data processing is likely to continue to accelerate with the digitization of most aspects of society and the expected adoption of virtual and augmented reality 'metaverses'.

Against this backdrop, timely technical, ethical, and policy questions arise: are biometric techniques designed to infer individuals' mental states scientifically sound? What are the individual and societal impacts of using these techniques to judge a person's qualities? What are the risks associated with collecting vast amounts of sensitive biometric data with the express intention of monitoring others' inner lives? How can and should such biometric techniques be regulated?

In the following sections we review use cases of biometric techniques that attempt to infer mental states and discuss their technical and ethical challenges. We then review the current regulatory landscape and close by presenting several considerations for the future regulation of these techniques.

Tracking your stress, fatigue, mental health... and everybody else's

Biometrics for personal monitoring Smartwatches and other wearable fitness trackers are increasingly popular, with at least one in five adults wearing one regularly in the United States [3]. Beyond counting steps, a variety of these devices claim to detect users' stress, mood, and fatigue by using various sensors and proprietary algorithms to collect and process continuously sampled data on movement, temperature, heart rate, voice patterns, and even brain activity. These devices and their associated smartphone applications notify users of their alleged mental states and prompt them to relax and focus through self-directed activity (e.g., via suggestions to take deep breaths, go for a walk, or change their tone of voice) or through soothing vibrations and nature sounds automatically initiated by the device.

For example, the Muse headband (by Canadian firm Interaxon) claims to translate brain activity captured from EEG electrodes on the forehead and behind the ears into "the guiding sounds of weather" to nudge users to focus and relax in real time. Similarly, Cove (by American firm Feelmore Labs) is worn behind the ears and produces vibrations to supposedly activate a brain pathway that helps users go from feeling tired and stressed out, to calm and "emotionally balanced."

Other, wrist-worn fitness trackers claim to detect stress and emotional states via data streams like electrodermal activity (EDA; variation in the skin's electrical properties in response to sweat secretion), heart rate (HR), and voice. For example, various wristbands from Fitbit (owned by Google) measure EDA and HR to infer a user's "stress response."

The device combines these data with exercise and sleep data to produce a daily "stress management score." Through the device's smartphone app, users can visualize their scores over time and take guided mindfulness and breathing sessions to improve their score.

As another example, the original version of Amazon's Halo wristband analyzes the user's voice data to identify distinct changes in pitch, pace, volume, and pauses and subsequently categorize their emotional tone, with labels from "annoyed", "excited", and "sad", to "curious", "hesitant" and "stubborn." Users can check how they sound in real-time and allow the device to listen continuously throughout the day to access a daily overview of their tone, including specific sentences the user said that illustrate each 'detected' tone.

Wearable biometric devices like these headbands and wristbands offer consumers the appeal of a healthier lifestyle with the assistance of cutting-edge, AI-powered technology. Through a sleek device and accompanying user-interface, users can keep closer track of their 'detected' mood and behavior via detailed data dashboards, and get support and motivation to improve their health and emotional wellbeing from the devices' various behavioral nudges.

However, questions remain regarding the scientific accuracy of their claims (see "Overstated scientific evidence and other technical challenges" below), and about how companies protect users' highly sensitive biometric data and related inferences from breaches and misuse. Privacy and security policies offered by fitness trackers and other wearables vary (e.g., only some de-identify the user data they collect), though most claim they will not sell users' data to third parties. Even with solid security protections from these companies, any data that users sync or share with third parties might be vulnerable to attacks. For example, in 2021 the health data of over 61 million Fitbit and Apple Watch users was exposed online by a third-party health and wellness company due to inadequate security protections [4].

Companies themselves might exploit their privileged access to the 'inferred' mental state of their users. For example, although supposedly not yet deployed in its consumer products, Amazon has patented the use of voice analyses to make its virtual assistant, Alexa, responsive to the users' 'perceived' physical and emotional state and to their "predicted intent, needs or desires" [5]. Alexa might offer to order cough drops when the user has a raspy voice, or serve music and advertisements based on whether the user "sounds happy, sad, or tired." While such features could advance the user experience by providing more relevant recommendations, they could also be exploited to manipulate users into adopting certain behaviors, including the purchasing of specific products or services [6]. This and other ethical issues will be covered in more detail below, under "Intensifying the ethical concerns of AI."

Biometrics for clinical research and monitoring Clinical investigators and practitioners are leveraging biometric techniques, such as 'digital phenotyping,' to conduct precision approaches to mental health research and care. Digital phenotyping researchers collect consenting participants' smartphone and wearable data (e.g., GPS, activity, heart rate, call and text logs, and brief surveys) to extract features (e.g., location type, sleep duration) that can be used to estimate mental and behavioral states of clinical relevance (e.g., social avoidance, insomnia).

A primary goal behind the use of behavioral biometrics in psychiatric research is to develop a more comprehensive characterization of the various physiological, behavioral and affective changes associated with conditions such as depression, bipolar disorder, schizophrenia, and suicidal ideation. The focus is not on attempting to 'read patients' minds' (indeed, participants are typically asked to self-report their current thoughts and mood), but rather to identify and characterize the correlates of more generalized states of distress and/or dysfunction. This approach seeks to move the field away from coarse, retrospective assessments done in the clinician's office, and towards the development of temporally precise disease phenotypes and markers. Ultimately, this would allow clinicians to better understand, diagnose, monitor, and treat clinical risk and disease, at scale and with relatively low burden to participants [7, 8].

Although digital phenotyping is a nascent methodology, recent peer-reviewed studies that analyzed patients' behavioral biometrics and brief surveys collected through wearables and smartphones show promise in identifying clinical risk states. For example, a study found that significant deviations from a patient's baseline behavior and self-reported emotional states, as captured through their smartphones, was indicative of heightened risk of schizophrenia relapse [9]. Another recent study found that surveys collected on patients' smartphones during psychiatric hospitalization outperformed a traditional clinical assessment in predicting suicide attempts in the month after discharge [10]. Such findings highlight the value of monitoring patients continuously over extended periods of time, in contrast to traditional clinical evaluations done at distant time points. Funding agencies like the National Institute of Mental Health have been signaling support for this research through substantial grants.

Interest in using biometric techniques to understand and treat mental health is also growing within the private sector, with digital health startups raising $5.1 billion in 2021 alone [11]. Increasingly, academics and private companies are collaborating through formal clinical research programs, like the partnership between Apple and the University of California Los Angeles to leverage iPhone and Apple Watch sensors for the monitoring of depression and cognitive decline [12]. Similarly, Verily Life Sciences, Alphabet Inc.'s research organization and an active collaborator with multiple universities, is developing a mobile app to collect sensor data to help identify risk for depression [13].

Within academic contexts, digital phenotyping research is highly regulated by institutional ethical review boards. These boards require researchers to conduct comprehensive

informed consent procedures to ensure study participants understand the extent of the data collected, the steps taken to protect their privacy and confidentiality, and how they can withdraw from the study and/or request the deletion of their data. Additionally, groups of academics have acknowledged the ethical challenges of using sensitive biometric and other personal data and, although optional, some have designed concrete guidelines for self-regulation [14]. There is currently considerably less transparency, scrutiny, and regulation of similar biometrics-based research that is conducted independently by private companies.

Biometrics for monitoring students, workers, and others

Biometric techniques that claim to infer mental states are not only deployed for personal tracking or the monitoring of patients within regulated clinical research contexts. They are also increasingly used in public and private settings to quantitatively judge individuals' intentions, engagement, and abilities, which subsequently informs important decisions such as border entry and educational and work opportunities.

Schools in China have deployed cameras in the classroom to monitor students' facial expressions and notify teachers and other authorities in real time if students were inattentive or had neutral and negative emotional expressions [15]. Chinese companies have also deployed ambient cameras and other sensors inserted in caps and office chairs to track their workers' facial expressions, brain activity, and posture in the attempt to infer their level of focus and emotional states like anxiety, rage, and sadness [16, 17].

The use of biometric techniques in the workplace is not limited to China. Several companies around the world, including in the U.S., now incorporate biometric techniques in their hiring practices, deploying software to analyze a candidate's voice intonation and facial expressions during an interview and produce various 'employability scores' that reflect their predicted job performance and suitability [18].

Across use cases, biometric techniques offer the appearance of greater objectivity, efficiency, and accuracy compared to purely human-based assessments. However, the claimed inference of mental states still suffers from several technical limitations.

Overstated scientific evidence and other technical challenges

While the use of biometric techniques in clinical research is regulated through established mechanisms like ethics review boards and peer-review, commercial products that use these techniques for the purported inference of mental states can be deployed without receiving much pre-market scrutiny over the accuracy of their claims, as long as they meet basic safety standards and do not make explicit medical claims regarding the diagnosis, prevention, or treatment of disease (1)(2) (see "Current legislation..." below).

In fact, several commercial products that supposedly detect or modify mental states lack rigorous scientific evidence. Some 'smart' headbands, like the ones covered above, allegedly activate specific brain pathways to provide a sense of calm, but their documentation is limited to in-house research reports that lack methodological and statistical detail and have not been peer-reviewed [19]. Similarly, several companies claim to accurately detect emotions from facial expressions, but academic researchers have repeatedly questioned these claims [20]–[22], pointing out that how people communicate emotions varies substantially across people, situations, and even within a single situation, and that unique emotion categories (e.g., fear, anger, sadness) do not necessarily have unique physiological signatures.

Even when biometric techniques are employed within rigorous academic research settings, substantial methodological challenges remain related to data quality and analysis. First, the impact of data missingness on the ability to estimate clinical risk states is understudied and could influence results. For example, a recent digital phenotyping study on participants with schizophrenia found that the amount of missing accelerometer and GPS data was significantly associated with the participants' symptom severity, suggesting that data missingness could be an important confounder that researchers should examine and account for in their analyses [23].

Other challenges in this research relate to validity and reliability. Current practices for validation often include comparing results of a new device to well-established laboratory instruments. For instance, one research team tried to estimate the error term of mobile heart rate sensors (e.g., in the Apple Watch) by having participants wear the device while connected to an electrocardiogram, a clinical-grade gold standard instrument [24]. However, there is no set standard on how much error is acceptable for new devices, so researchers have little guidance on how much error to tolerate, while private companies are given plenty of room to claim their products are validated [25].

Finally, ensuring reliability is also difficult. One researcher working with Apple Watches recently found that downloading data from the same watch at two different timepoints did not yield the same dataset [26]. Apple had likely updated their underlying data processing algorithms, but the researcher was unaware of when or why this change had occurred. This highlights the tension between needing to collect reliable biometric data over extended periods of time (e.g., to monitor clinical symptom progression) and the fact that technologies are constantly updated.

Intensifying the ethical concerns of AI

Although biometric techniques that attempt to infer mental states contain multiple technical challenges, some of them are already deployed in commercial products that present to the user as though they are providing accurate readings of emotions, intentions, and skills. Such uses of these techniques, especially when outside of regulated and transparent research settings, magnify the ethical risks of artificial intelligence applications concerning people's rights, choices, and well-being [27]. Among others, these risks

encompass the widening scope of real-life decisions subjected to biased algorithmic profiling, unprecedented levels of surveillance of public and private life, and the increasing scale of sensitive data profiles vulnerable to breach and exploitation. (For more extensive critical analysis on the various societal impacts of using technology to quantify bodily and mental states, see work by scholars in anthropology and science and technology studies such as [28] and [29]). Together, these dimensions weave an uncharted territory of moral hazard that warrants special attention from policymakers.

Algorithmic profiling of inner features with lasting consequences The widening private and public settings deploying biometric techniques for the purported inference of others' mental states implies that people might no longer be assessed by their deliberate actions, but also by their involuntary physiological signals or 'inferred' intentions that they have not acted upon. These dubious inferences and categorizations of individuals' mental states are particularly risky when they are used to make consequential decisions for the data subjects across many aspects of their life, such as facial analysis software that ranks a job applicant's employability or smart border systems that flag the 'inferred' aggression level of an individual for further examination. If data operators believe these algorithmic inferences to offer holistic and decisive profiles of the data subjects, they might lead to persistent differential treatment of individuals over time. For instance, so-called emotion recognition systems that categorize students' mental states based on their eye movements and posture could shape how teachers evaluate their students' academic potential over months or even years.

Although such uses of biometric techniques are unreliable across the board, their negative consequences might disproportionally affect certain groups based on their race, age, gender, cultural expression, and behavioral characteristics [30]. Gait, facial expression, voice and other biometrics are uniquely personal, but may vary by demographic group. The training of these AI models on historical biometric data, particularly when certain groups are under-represented, will likely lead to results that are biased, which then reinforce discriminatory feedback loops for future data subjects. For example, certain emotional AI applications already interpret black faces as "having more negative emotions" than white faces [31], and supposedly perceive the emotions of younger adults more "accurately" than those of elderly people [32].

Unprecedentedly invasive surveillance and violation of privacy and autonomy Biometric technologies attempting to detect mental states could escalate from relatively deliberate data gathering for targeted scrutiny to a form of total, always-on surveillance [33]. From sensors embedded in wearables and home devices to cameras sampling micro-expressions in classrooms and the workplace, this elevated form of surveillance threatens to leave no individual or no particular moment unexamined, and thus to severely undermine individuals' right to privacy.

Even if the inferences made by these techniques are known to be faulty, the systematic scrutiny of sensitive biometric data and ordinary activities disempowers individuals from behaving in an autonomous and uninhibited way, and can lead to hypervigilance and distress [28, 34]. Indeed, these concerns are likely exacerbated by knowing that these technologies are flawed and might lead to unfair treatment.

Fine-grained biometrics at risk of data breaches and exploitation Biometric techniques share the general cybersecurity vulnerabilities of other digital technologies related to the unauthorized collection, access, and use of data by third parties or by trusted operators themselves [35]. However, biometric data systematically collected to estimate mental states could be prone to additional risks of exploitation and misuse due to the volume and granularity of the data involved. Despite the scientific flaws in mental state inference, the perceived value in stealing fine-grained, continuous behavioral biometric data that might map people's emotions and intent could attract new forms of malicious attacks for commercial or political gains. The trend of 'mood marketing' already hints at such appeal even when biometrics are not involved, with companies like Spotify and Facebook using clicks, views, and keystrokes to infer users' moods in real-time and curate targeted information [36]. This is particularly concerning considering biometric data obtained through wearables and cameras can be difficult for people to modify or opt out of and that many devices lack sufficient processing power to deploy strong security mechanisms [37].

Mapping further impacts: from accountability to long-term consequences The risks presented above are not exhaustive. A challenge in evaluating biometric technologies for mental inference is that more abstract ethical impacts might be harder to measure and implement relative to technical fixes for algorithmic bias or security protocols [38]. For instance, what if presenting a user with frequent notifications of 'detected' negative mood triggers an emotional negativity tailspin? To what extent should these technologies bear accountability for such emotional experiences [29, 39]? Longer-term social impacts, such as the influence of alleged mood-detection technology on children and adolescents' emotional development, might also take root without immediate policy attention.

As the market of biometric technologies rapidly expands and transforms intimate aspects of the human experience, policymakers are tasked with designing regulatory frameworks that carefully consider their associated harms and risks.

Current legislation and general considerations for regulation

The regulation of biometric techniques that claim to detect mental states can take several forms, including legislation, guidelines by international and nongovernmental organizations, institutional review boards in research settings, industry standards, and educational initiatives to raise public awareness and promote digital skills. Among them, legislation is likely to have the largest impact.

As with pharmacology and genetics, innovations in biometric technologies may bring benefits and harms to individuals and the public. The challenge for policymakers is to craft legislative solutions that minimize risk without amplifying other harms or severely obstructing technological innovation and its potential benefits in fields like clinical research and healthcare, as discussed earlier.

While the European Union is the most advanced in regulating the use of biometric data for purposes beyond the identification of individuals, regulation in Europe is still sparse and mostly at the proposal stage. Federal legislation in the United States protects biometric and health data only in certain contexts, and the few states that offer wider biometric data protections have focused on biometrics-based identification without addressing the attempted inference of mental states. General considerations for the regulation of biometric techniques moving forward are presented at the end of this section.

Europe The EU has taken the lead in creating a comprehensive legislative framework for the regulation of various emerging issues related to the digital revolution, including biometric technologies. One of the first such initiatives was the **General Data Protection Regulation (GDPR)** (3), enforced in May 2018. The GDPR has led to vast improvements in awareness and standards of privacy and personal data protection, including the establishment of overarching principles such as fairness and transparency, consent, data minimization, purpose limitation, data protection by design and by default, and a risk-based approach, which sets requirements for legitimate data processing in proportion to the risks posed to data subjects. Scientific research is considered a special category under the GDPR that receives specific exemptions related to the processing of biometrics and other personal health data, but always with appropriate safeguards and subject to ethical standards and most data protection conditions mentioned above.

Nevertheless, the GDPR has been criticized for lacking effective enforcement mechanisms, its unrealistic approach to legitimate consent, and its insufficient regulation of algorithmic data processing for profiling and automated decision-making [40, 41]. Regarding biometric techniques in particular, the GDPR introduced a legal definition of biometric data as "personal data resulting from specific technical processing relating to the physical, physiologic or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data". Although imposing elevated requirements for the processing of these data, the focus is on first-generation biometrics used exclusively for the *identification* of individuals, thus posing no legal requirements for the use of behavioral biometrics for the assessment of individuals' *mental states*.

Another important milestone in the regulation of biometric techniques is the draft **Artificial Intelligence Act (AIA)** (4). The AIA is a pioneering attempt at horizontal regulation of the development, commodification, and use of AI applications, including biometric techniques. While adopting the GDPR's definition of biometric data, the draft AIA introduces legal definitions and risk-based requirements regarding biometric data processing in so-called emotion recognition systems and biometric categorization systems. The proposal introduces four levels of risk, from minimal to unacceptable. Biometric data processing is qualified as a high-risk AI system, meaning it is not entirely prohibited but subject to additional requirements such as ensuring transparency, conformity with necessary data protection and assessment procedures, and implementing appropriate risk management and human oversight. The only type of AI system qualified as posing unacceptable risk is real-time remote biometric *identification* systems in public places, e.g., via facial recognition.

Although the draft AIA provides a more granular differentiation of requirements for technologies according to their risk level and prohibits certain uses of AI altogether, the AIA is criticized for sticking to the narrow definition of biometric data provided in the GDPR, which leaves out the use of behavioral biometrics for the attempted inference of individuals' *mental states*. A range of scholars [42], international non-governmental organizations [43], and European data protection authorities [44] are advocating for the AIA to prohibit the use of 'emotion detection' AI systems (particularly those based on facial expression data) altogether, not only because of the individual and societal risks they pose, but also due to their pseudoscientific basis.

United States Federal data protection regulations are still scarce in the United States. The **Health Insurance Portability Accountability Act (HIPAA)** (5) regulates the use and sharing of individuals' personal health information (PHI), including many biometric data streams mentioned in this article. However, HIPAA only covers PHI handled by healthcare providers, health insurance companies, and their business associates. Thus, many of the consumer products covered in this paper, from personal fitness trackers to the sensors deployed in workplace and educational settings, are not subject to HIPAA regulation, even though they collect health information and/or biometric data that could be used to assess an individual's (physical or mental) health status.

Similarly, the **Federal Food, Drug and Cosmetics Act**, enforced by the U.S. Food and Drug Administration (FDA) (1), only regulates products that make explicit claims regarding the diagnosis, treatment, or prevention of a disease or condition. Products that are broadly "intended for maintaining or encouraging a healthy lifestyle" in the absence of explicit medical claims, including fitness trackers and applications that claim to infer users' stress levels and mood, are typified as "general wellness products" that pose low risk to the users' safety. As such, these products are not subject to the FDA's pre-market review and post-market regulations pertaining quality control, labeling, and periodic evaluation and reporting requirements.

In terms of general data protection, the U.S. Federal Trade

Commission (FTC) (6), under Section 5 of the **FTC Act**, protects against "unfair or deceptive trade practices" and has been consequential in enforcing hundreds of cases in which companies do not comply with their own data privacy and data security policies. However, the FTC's authority is limited: it does not prescribe specific policies and does not regulate companies that have not made explicit promises regarding data protection.

A few state laws, such as **Illinois' Biometric Information Privacy Act** (7), and the proposed National Biometric Information Privacy Act of 2020 (8) demonstrate an interest in regulating biometric data processing more directly (including requiring written consent from the data subjects) and might serve as a foundation for more comprehensive regulation. However, similar to the EU's GDPR, existing legislation and proposals alike have focused exclusively on traditional biometrics (e.g., iris scans, fingerprints, face geometry) used for the identification and authentication of individuals. The **California Consumer Privacy Act** (9), modeled after Europe's GDPR, provides the broadest definition of biometric data of all other U.S. legislations and proposals, including behavioral biometrics, but still only protects against its use for the purposes of identification, and makes no mention of the attempted inference of mental states.

China Enforced in November 2021 and largely inspired by the GDPR, China's **Personal Information Protection Law (PIPL)** (10) classifies biometric information as sensitive data whose processing is subject to additional requirements and protections, including notifying the data subjects of the necessity and consequences of such actions. Similar to the purpose limitation principle in the European legislation, the PIPL dictates that data processing may be performed only for specific purposes, but these provisions are as broad as in the GDPR, with a focus on traditional biometric data used for the identification of individuals.

Current legislation and general considerations for regulation

Given the wide range of data streams processed, the dubious scientific grounds of many commercial products, their diverse use cases, and the yet undetermined reach of their societal impacts, the regulation of biometric techniques that attempt to infer mental states is undoubtedly difficult and complex. Below we highlight a few general considerations for the regulation of these techniques worldwide.

Clarifying the scope of regulation Most of the existing legislation in this space defines biometric data only in relation to facial recognition and other first-generation biometrics used for the *identification* of individuals. This excludes other risky forms of physical, physiological, or behavioral data processing from being regulated, including the use of voice, facial expression, heart rate, and brain activity to attempt to infer mental states. Expanding and clarifying the legal definitions of biometric data and biometric techniques would close the legal loopholes that currently allow risky technologies to escape legal requirements and accountability [43].

Building awareness and transparency around all uses of biometric techniques As reviewed extensively in this review, the collection and processing of biometric data poses several risks to individuals and society, even in the absence of identification. Building the public's long-term trust in biometric techniques more broadly will likely require, among others, full transparency regarding their scientific evidence (or lack thereof), the disclosure of all current uses of biometric techniques and their known and potential risks, clear privacy and security protections, and the respect for individuals' autonomy via informed consent and opt-in (rather than opt-out) procedures by default.

Balancing proactive and reactive approaches To enhance biometric technologies' quality and human-centric nature in the long term, legislation could not only concern the regulation of techniques once they are in use, but also provide proactive measures enabling authorities to monitor the entire product lifecycle. Monitoring during the early phases of product design and testing is vital to assess the validity and reliability of the methods used and curtail potential individual and societal harms from the beginning. Instruments such as the GDPR's data protection impact assessment or the human rights impact assessment lobbied by human rights organizations [43], together with adequate market entrance requirements such as the verification of the product's claims, might be equally or more important as post-market surveillance processes. Simultaneously, regulation should be flexible enough to allow for future amendments and updates that keep pace with rapidly evolving technologies and uses of them.

Incorporating effective enforcement and redress mechanisms To avoid some of the enforcement problems faced by Europe's GDPR, it will be important to thoroughly plan the monitoring procedures and furnish the responsible enforcement authorities with sufficient competences, resources, and power (e.g., to award sufficiently deterrent fines for noncompliance). Regulations might also consider facilitating complaints from individuals or groups impacted by biometric data processing and provide them with access to proportionate remedies.

International cooperation for harmonized and sustainable regulation As new regulations on emerging technologies arise in different regions of the world, the international nature of technology markets and the massive scale of multinational data flows means that maintaining incompatible legal requirements will likely lead to significant disruptions. The sustainable development and regulation of biometric technologies may therefore require systematic multilateral cooperation to harmonize legal requirements across jurisdictions and produce internationally accepted standards and protocols for the protection of human rights [45].

Conclusions and the path forward

Technologies that use behavioral biometric data to attempt to infer mental states—from momentary emotions to more stable intentions, abilities, and health status—promise greater efficiency and accuracy in the assessment of individuals'

minds in contexts like healthcare, education, work, and law enforcement. However, many of these techniques, especially those deployed outside of rigorous and regulated research settings, lack adequate scientific evidence and present serious technical and ethical issues. Critically, current uses of commercialized biometric techniques are already making consequential decisions about individuals' emotions, intentions and skills based on faulty inferences, attempting to surveil individuals' public and private lives, and exposing sensitive biometric data to breaches and exploitation.

A better understanding of these technologies and their consequences is critical to build regulatory frameworks that mitigate harms while allowing for biometric techniques' potential benefits, such as for consensual and regulated clinical monitoring and care. While this review contributed to this goal, additional research is needed to more comprehensively assess the limits of what behavioral biometrics can reveal about mental states, the benefits of leveraging biometric techniques in healthcare and beyond, and the various short- and long-term impacts to individuals, groups, and society at large.

In addition to rigorous research, the successful regulation of biometric techniques for the attempted inference of mental states will likely require more precise legal definitions of biometric data and their uses, pre- and post-market evaluations of these techniques' claims and risks, and regulatory harmonization across jurisdictions.

Citation

Vidal Bustamante, C. M., Alama-Maruta, K., Ng, C., Coppersmith, D. D. L. Should machines be allowed to 'read our minds'? Uses and regulation of biometric techniques that attempt to infer mental states. *MIT Science Policy Review* **3**, 112-121 (2022). https://doi.org/10.38105/spr.qy2iibrk72.

Open Access

Legislation Cited

(1) U.S. Food and Drug Administration's General Wellness Products Draft Guidance for Industry and Food and Drug, Policy for Low Risk Devices, FDA-2014-N-1039.
(2) Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices.
(3) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
(4) Proposal for a Regulation of the European Parliament and of the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts COM/2021/206 final.
(5) Health Insurance Portability and Accountability Act of 1996. Pub. L. No. 104-191, § 264, 110 Stat.1936.
(6) Federal Trade Commission Act. 15 U.S.C. §§ 41-58, as amended.
(7) Biometric Information Privacy Act, 740 ILCS 14/.
(8) National Biometric Information Privacy Act, S. 4400, 116th Congress (2020).
(9) California Consumer Privacy Act, 2018 California Legislative Service Ch. 55 (A.B. 375).
(10) China's Personal Information Protection Law. Official translation provided by China's National People's Congress available at http://en.npc.gov.cn.cdurl.cn/2021-12/29/c_694559.htm

References

[1] Yannopoulos, A., Andronikou, V. & Varvarigou, T. Behavioural biometric profiling and ambient intelligence. In *Profiling the European Citizen*, 89–109 (Springer, 2008). https://doi.org/10.1007/978-1-4020-6914-7_5.
[2] Wendehorst, C. & Duller, Y. Biometric recognition and behavioural detection: Assessing the ethical aspects of biometric recognition and behavioural detection techniques with a focus on their current and future use in public spaces (2021). Online: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/696968/IPOL_STU(2021)696968_EN.pdf.
[3] Vogels, E. About one-in-five americans use a smart watch or fitness tracker. *Pew Research Center* (2020). Online: https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/.
[4] McKeon, J. 61M Fitbit, Apple users had data exposed in wearable device data breach. *Patient Privacy News, Health IT Security* (2021). Online: https://healthitsecurity.com/news/61m-fitbit-apple-users-had-data-exposed-in-wearable-device-data-breach.
[5] Jin, H. & Wang, S. Patent US10096319b1: Voice-based determination of physical and emotional characteristics of users. *Amazon* (2017). Online: https://patents.google.com/patent/US10096319B1/en.
[6] Turow, J. *The Voice Catchers* (Yale University Press, 2021).
[7] Insel, T. R. Digital phenotyping: technology for a new science of behavior. *JAMA* **318**, 1215–1216 (2017). https://doi.org/10.1001/jama.2017.11295.
[8] Onnela, J.-P. & Rauch, S. L. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology* **41**, 1691–1696 (2016). https://doi.org/10.1038/npp.2016.7.
[9] Henson, P., D'Mello, R., Vaidyam, A., Keshavan, M. & Torous, J. Anomaly detection to predict relapse risk in schizophrenia. *Translational Psychiatry* **11**, 1–6 (2021). https://doi.org/10.1038/s41398-020-01123-7.
[10] Wang, S. B. *et al.* A pilot study using frequent inpatient assessments of suicidal thinking to predict short-term postdischarge suicidal behavior. *JAMA Network Open* **4**, e210591–e210591 (2021). https://doi.org/10.1001/jamanetworkopen.2021.0591.

[11] Krasniansky, A., Evans, B. & Zweig, M. 2021 year-end digital health funding: Seismic shifts beneath the surface. *Rock Health* (2022). Online: https://rockhealth.com/insights/2021-year-end-digital-health-funding-seismic-shifts-beneath-the-surface/.

[12] Winkler, R. Apple is working on iPhone features to help detect depression, cognitive decline. *Wall Street Journal* (2021). Online: https://www.wsj.com/articles/apple-wants-iphones-to-help-detect-depression-cognitive-decline-sources-say-11632216601?mod=rss_Technology.

[13] Nickels, S. *et al.* Toward a mobile platform for real-world digital measurement of depression: User-centered design, data quality, and behavioral and clinical modeling. *JMIR Mental Health* **8**, e27589 (2021). https://doi.org/10.2196/27589.

[14] Shen, F. X. *et al.* An ethics checklist for digital health research in psychiatry. *Journal of Medical Internet Research* **24**, e31146 (2022). https://doi.org/10.2196/31146.

[15] Marda, V. & Ahmed, S. Emotional entanglement: China's emotion recognition market and its implications for human rights. *Article 19* (2021). Online: https://www.article19.org/wp-content/uploads/2021/01/ER-Tech-China-Report.pdf.

[16] Chen, S. Forget the facebook leak: China is mining data directly from workers' brains on an industrial scale. *South China Morning Post* (2018). Online: https://www.scmp.com/news/china/society/article/2143899/forget-facebook-leak-china-mining-data-directly-workers-brains.

[17] Sun, N. China's tech workers pushed to limits by surveillance software. *Nikkei Asia* Online: https://asia.nikkei.com/Spotlight/The-Big-Story/China-s-tech-workers-pushed-to-limits-by-surveillance-software.

[18] Sloane, M., Moss, E. & Chowdhury, R. A silicon valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns* **3**, 100425 (2022). https://doi.org/10.1016/j.patter.2021.100425.

[19] Study: EEG Imaging Reveals that Cove Triggers Relaxation by Activating a Powerful Brain Pathway. *Cove* (2017). Online: https://www.feelcove.com/2021-04-01/Study_EEG_Summary.pdf.

[20] Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M. & Pollak, S. D. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest* **20**, 1–68 (2019). https://doi.org/10.1177/1529100619832930.

[21] Siegel, E. H. *et al.* Emotion fingerprints or emotion populations? a meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin* **144**, 343 (2018). https://doi.org/10.1037/bul0000128.

[22] Leys, R. *The Ascent of Affect: Genealogy and Critique* (University of Chicago Press, 2017).

[23] Torous, J. *et al.* Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. *NPJ Digital Medicine* **1**, 1–9 (2018). https://doi.org/10.1038/s41746-018-0022-8.

[24] Bent, B., Goldstein, B. A., Kibbe, W. A. & Dunn, J. P. Investigating sources of inaccuracy in wearable optical heart rate sensors. *NPJ Digital Medicine* **3**, 1–9 (2020). https://doi.org/10.1038/s41746-020-0226-6.

[25] Huckvale, K., Venkatesh, S. & Christensen, H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digital Medicine* **2**, 1–11 (2019). https://doi.org/10.1038/s41746-019-0166-1.

[26] Onnela, J.-P. Exporting the same data from a wearable twice doesn't give you the same data. *Beiwe* (2021). Online: https://www.beiwe.org/exporting-the-same-data-from-a-wearable-twice-doesnt-give-you-the-same-data/.

[27] Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. & Srikumar, M. Principled artificial intelligence: Mapping consensus in ethical and rights-based approaches to principles for ai. *Berkman Klein Center Research Publication* (2020). http://dx.doi.org/10.2139/ssrn.3518482.

[28] Nafus, D. *Quantified: Biosensing technologies in everyday life* (MIT Press, 2016).

[29] Ruckenstein, M. & Schüll, N. D. The datafication of health. *Annual Review of Anthropology* **46**, 261–278 (2017). https://doi.org/10.1146/annurev-anthro-102116-041244.

[30] Leslie, D. Understanding bias in facial recognition technologies. *SSRN* **4050457** (2020). https://doi.org/10.48550/arXiv.2010.07023.

[31] Rhue, L. Racial influence on automated perceptions of emotions. *SSRN* **3281765** (2018). https://dx.doi.org/10.2139/ssrn.3281765.

[32] Kim, E., Bryant, D., Srikanth, D. & Howard, A. Age bias in emotion detection: an analysis of facial emotion recognition performance on young, middle-aged, and older adults. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 638–644 (2021). https://doi.org/10.1145/3461702.3462609.

[33] Lyon, D. Surveillance, power and everyday life. In *Emerging digital spaces in contemporary society*, 107–120 (Springer, 2010). https://doi.org/10.1057/9780230299047_18.

[34] Mantello, P., Ho, M.-T., Nguyen, M.-H. & Vuong, Q.-H. Bosses without a heart: socio-demographic and cross-cultural determinants of attitude toward Emotional AI in the workplace. *AI & society* 1–23 (2021). https://doi.org/10.1007/s00146-021-01290-1.

[35] How biometrics are attacked. *UK National Cyber Security Centre: Guidance on Biometric Recognition and Authentication Systems* (2019). Online: https://www.ncsc.gov.uk/collection/biometrics/how-biometrics-are-attacked.

[36] Mahdawi, A. Spotify can tell if you're sad. here's why that should scare you. *The Guardian* (2018). Online: https://www.theguardian.com/commentisfree/2018/sep/16/spotify-can-tell-if-youre-sad-heres-why-that-should-scare-you.

[37] Pal, S., Mukhopadhyay, S. & Suryadevara, N. Development and progress in sensors and technologies for human emotion recognition. *Sensors* **21**, 5554 (2021). https://doi.org/10.3390/s21165554.

[38] Whittaker, M. *et al.* *AI Now Report 2018* (AI Now Institute at New York University, New York, 2018). Online: https://ainowinstitute.org/AI_Now_2018_Report.pdf.

[39] Kaplan, M. Happy with a 20% chance of sadness. *Nature* **563**, 20–23 (2018). https://doi.org/10.1038/d41586-018-07181-8.

[40] Three years under the EU GDPR: An implementation progress report. *Access Now* (2021). Online: https://www.accessnow.org/cms/assets/uploads/2021/05/Three-Years-Under-GDPR-report.pdf.

[41] Brkan, M. Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the gdpr and beyond. *International Journal of Law and Information Technology* **27**, 91–121 (2019). https://doi.org/10.1093/ijlit/eay017.

[42] Veale, M. & Borgesius, F. Z. Demystifying the draft eu artificial intelligence act—analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International* **22**, 97–112 (2021). https://doi.org/10.9785/cri-2021-220402.

[43] Access Now's submission to the European Commission's adoption consultation on the Artificial Intelligence Act. *Access Now* (2021). Online: https://www.accessnow.org/cms/assets/uploads/2021/08/Submission-to-the-European-Commissions-Consultation-on-the-Artificial-Intelligence-Act.pdf.

[44] EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial

Intelligence Act). *European Data Protection Board and European Data Protection Supervisor* (2021). Online: https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en.

[45] Kop, M. EU Artificial Intelligence Act: The European Approach to AI (Transatlantic Antitrust and IPR Developments, 2021). Online: https://ssrn.com/abstract=3930959.