

World view



By Kate Crawford

Time to regulate AI that interprets human emotions

The pandemic is being used as a pretext to push unproven artificial-intelligence tools into workplaces and schools.

During the pandemic, technology companies have been pitching their emotion-recognition software for monitoring workers and even children remotely. Take, for example, a system named 4 Little Trees. Developed in Hong Kong, the program claims to assess children's emotions while they do classwork. It maps facial features to assign each pupil's emotional state into a category such as happiness, sadness, anger, disgust, surprise and fear. It also gauges 'motivation' and forecasts grades. Similar tools have been marketed to provide surveillance for remote workers. By one estimate, the emotion-recognition industry will grow to US\$37 billion by 2026.

There is deep scientific disagreement about whether AI can detect emotions. A 2019 review found no reliable evidence for it. "Tech companies may well be asking a question that is fundamentally wrong," the study concluded (L. F. Barrett *et al. Psychol. Sci. Public Interest* 20, 1–68; 2019).

And there is growing scientific concern about the use and misuse of these technologies. Last year, Rosalind Picard, who co-founded an artificial intelligence (AI) start-up called Affectiva in Boston and heads the Affective Computing Research Group at the Massachusetts Institute of Technology in Cambridge, said she supports regulation. Scholars have called for mandatory, rigorous auditing of all AI technologies used in hiring, along with public disclosure of the findings. In March, a citizen's panel convened by the Ada Lovelace Institute in London said that an independent, legal body should oversee development and implementation of biometric technologies (see go.nature.com/3cejmktk). Such oversight is essential to defend against systems driven by what I call the phrenological impulse: drawing faulty assumptions about internal states and capabilities from external appearances, with the aim of extracting more about a person than they choose to reveal.

Countries around the world have regulations to enforce scientific rigour in developing medicines that treat the body. Tools that make claims about our minds should be afforded at least the same protection. For years, scholars have called for federal entities to regulate robotics and facial recognition; that should extend to emotion recognition, too. It is time for national regulatory agencies to guard against unproven applications, especially those targeting children and other vulnerable populations.

Lessons from clinical trials show why regulation is important. Federal requirements and subsequent advocacy have made many more clinical-trial data available

There is deep scientific disagreement about whether AI can detect emotions."

Kate Crawford is a research professor at University of Southern California Annenberg in Los Angeles and a senior principal researcher at Microsoft Research in New York. Her latest book is *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*.
Twitter: @katecrawford

to the public and subject to rigorous verification. This becomes the bedrock for better policymaking and public trust. Regulatory oversight of affective technologies would bring similar benefits and accountability. It could also help in establishing norms to counter over-reach by corporations and governments.

The polygraph is a useful parallel. This 'lie detector' test was invented in the 1920s and used by the FBI and US military for decades, with inconsistent results that harmed thousands of people until its use was largely prohibited by federal law. It wasn't until 1998 that the US Supreme Court concluded that "there was simply no consensus that polygraph evidence is reliable".

A formative figure behind the claim that there are universal facial expressions of emotion is the psychologist Paul Ekman. In the 1960s, he travelled the highlands of Papua New Guinea to test his controversial hypothesis that all humans exhibit a small number of 'universal' emotions that are innate, cross-cultural and consistent. Early on, anthropologist Margaret Mead disputed this idea, saying that it discounted context, culture and social factors.

But the six emotions Ekman described fit perfectly into the model of the emerging field of computer vision. As I write in my 2021 book *Atlas of AI*, his theory was adopted because it fit what the tools could do. Six consistent emotions could be standardized and automated at scale – as long as the more complex issues were ignored. Ekman sold his system to the US Transportation Security Administration after the 11 September 2001 terrorist attacks, to assess which airline passengers were showing fear or stress, and so might be terrorists. It was strongly criticized for lacking credibility and for being racially biased. However, many of today's tools, such as 4 Little Trees, are based on Ekman's six-emotion categorization. (Ekman maintains that faces do convey universal emotions, but says he's seen no evidence that automated technologies work.)

Yet companies continue to sell software that will affect people's opportunities without clearly documented, independently audited evidence of effectiveness. Job applicants are being judged unfairly because their facial expressions or vocal tones don't match those of employees; students are being flagged at school because their faces seem angry. Researchers have also shown that facial-recognition software interprets Black faces as having more negative emotions than white faces do.

We can no longer allow emotion-recognition technologies to go unregulated. It is time for legislative protection from unproven uses of these tools in all domains – education, health care, employment and criminal justice. These safeguards will recentre rigorous science and reject the mythology that internal states are just another data set that can be scraped from our faces.

CATHMUSCAT