

## **STE Episode Transcript: Affectiva**

**RANA EL KALIOUBY:** A lip corner pull, which is pulling your lips outwards and upwards – what we call a smile – is action unit 12.

The face is one of the most powerful ways of communicating human emotions.

We're on this mission to humanize technology.

I was spending more time with my computer than I did with any other human being – and this computer, it knew a lot of things about me, but it didn't have any idea of how I felt.

There are hundreds of different types of smiles. There's a sad smile, there's a happy smile, there's an "I'm embarrassed" smile – and there's definitely an "I'm flirting" smile.

When we're face to face, when we actually are able to get these nonverbal signals, we end up being kinder people.

The algorithm tracks your different facial expressions.

I really believe this is going to become ubiquitous. It's going to be the standard human machine interface.

Any technology in human history is neutral. It's how we decide to use it.

There's a lot of potential where this thing knows you so well, it can help you become a better version of you. But that same data in the wrong hands could be manipulated to exploit you.

**CATERINA FAKE:** That was computer scientist Rana el Kaliouby. She invented an AI tool that can read the expression on your face and know how you're feeling in real time.

With this software, computers can read our signs of emotion – happiness, fear, confusion, grief – which paves the way for a future where technology is more human, and therefore serves us better. But in the wrong hands, this emotion-reading engine could take advantage of us at our most vulnerable moments.

Rana is on the show because she wants to guide her technology towards a future that benefits all of us.

**[THEME MUSIC]**

**FAKE:** I'm Caterina Fake, and I believe that the boldest new technologies can help us flourish as human beings. Or destroy the very thing that makes us human. It's our choice.

A bit about me: I co-founded Flickr – and helped build companies like Etsy, Kickstarter and Public Goods from when they first started. I'm now an investor at Yes VC, and your host on *Should This Exist?*.

On today's show: An AI tool that can read the expression on your face – or the tone in your voice – and know how you're feeling. This AI tool picks up on your non-verbal cues – from the furrow in your brow to the pitch of your voice – to know whether you're happy or sad, surprised or angry.

The company is called Affectiva and they've trained their system using four billion facial frames, all from people who've given their consent.

This technology has soaring potential and dark shadows. On the one hand, we stand to gain so much by bringing human emotion into technology. A computer that knows when you're confused – or bored – can help you learn. A car that knows when you're distracted can keep you safe. A phone that knows you're sad can have your friends text you the right thing at the right time.

But let's be real: In the wrong hands, this technology is terrifying. Any system that knows how you feel could be used to influence or incriminate you. Think about "thoughtcrime" in *1984*. Or "precrime" in the "Minority Report".

This is an invention that sparks a lot of skepticism in a lot of people for a lot of different reasons.

**ESTHER PEREL:** It's not new that people want to try to capture your imagination, your emotional imagination, or any other. It's an old fantasy. It's an old dream.

**FAKE:** That was the renowned couples therapist Esther Perel. You might know her from her TED Talks or her podcast *Where Should We Begin?* We'll hear more from Esther and other voices later in the show, for our workshop. That's where we really dig into all sides of the question: Should this exist?

But I want to start with Affectiva's co-founder and CEO, Rana El-Kaliouby.

The first question I always ask about any invention is: Who created it? And what problem are they trying to solve?

**EL KALIOUBY:** I'm on a mission to humanize technology. I got started on this path over 20 years ago. My original idea was to improve human-computer interfaces and make those more compassionate and empathetic. But I had this "aha moment" that it's not just

about human-machine communications. It's actually fundamentally about how humans connect and communicate with one another.

**FAKE:** Rana believes that humans could communicate with one another better if our technologies could grasp human emotion, kind of the way humans do on our own. I was drawn in by her line of thinking.

**FAKE:** I'm sitting here with you and we're in person, and I can see the way your face moves, your micro expressions, moments of doubt, or thought, or amusement. And so we're designed for this. We're made for this.

**EL KALIOUBY:** Exactly, humans are hardwired for connection. That's how we build trust, that's how we communicate, that's how we get stuff done.

**FAKE:** And when you're communicating with somebody through chat, let's say, how much is lost?

**EL KALIOUBY:** Ninety three percent of our communication is nonverbal. Only 7% is communicated via the actual words we choose to use. So when you're texting somebody, you're just communicating 7% of the signal, and 93% is just lost.

**FAKE:** Right.

**EL KALIOUBY:** And that's what we're trying to fix. We're trying to redesign technology, to allow us to incorporate this 93% of our communication today.

**FAKE:** Mhmm.

**EL KALIOUBY:** I also feel when we're face to face, when we actually are able to get these nonverbal signals, we end up being kinder people.

**FAKE:** Well, it's kind of like road rage, right?

**EL KALIOUBY:** Exactly.

**FAKE:** Road rage comes from: There's not really a person in there, it's just a car that's in your way, or cut you off in traffic.

**FAKE:** Talking with Rana, I began thinking about the ways our cars – or computers or phones might serve us better if they felt more human. But there are many ways to humanize technology. And Rana has focused on interpreting nonverbal signals, like facial expressions.

And it turns out, she's been studying facial expressions a long time. Rana was born in Cairo, and moved to Kuwait as a kid.

**EL KALIOUBY:** My dad was pretty strict. I'm one of three daughters. And so he had very strict rules around what we could and could not do. And so one of those rules was we were not allowed to date. In fact, I was not allowed to date until I finished university. So my first date was right out of college.

And so as a result, I just basically focused on other kids at school and I became an observer, right? So I remember, especially as a high schooler, sitting during recess in our school's playground and just observing people. There was one particular couple, I think I spotted their flirting behavior even before they both realized it

The things that I remember clearly are the change in eye blinking behavior, the eyes lock for an extra few more seconds than they usually do. I became fascinated by all of these unspoken signals that people very subconsciously exchange with each other.

**FAKE:** Rana's fascination ultimately became the focus of her career. She studied computer science at the American University in Cairo, and pursued her PhD at Cambridge. She was newly married, but her husband was back home in Cairo. And late one night, in the Cambridge computer lab, she had an aha moment.

**EL KALIOUBY:** I realized very quickly that I was spending more time with my computer than I did with any other human being. This computer, despite the intimacy we had – it knew a lot of things about me: it knew what I was doing, it knew my location, it knew who I was – but it didn't have any idea of how I felt.

But even worse, it was my main portal of communication to my family back home. I often felt that all my emotions, the subtlety and the nuances of everything I felt disappeared in cyberspace.

There was one particular day, I was in the lab, it was later at night, and I was coding away. I was feeling really, really homesick. And I was in tears, but he couldn't see that. And I felt that technology is dehumanizing us, it's taking away of this fundamental way of how we connect and communicate.

**FAKE:** You were trying to reach people. You were trying to connect to people. You were far from everything that you'd ever known, right?

**EL KALIOUBY:** Right.

**FAKE:** And I think a lot of people, we're far from home.

**FAKE:** At Cambridge, Rana developed an emotionally intelligent machine – which recognized complex emotional states, based on facial expressions. She won a National Science Foundation grant to come to the U.S. and join the MIT Media Lab. There, Rana took this core technology and applied it to different products. The first use case: Kids with autism.

**EL KALIOUBY:** People on the autism spectrum really have trouble understanding and reading other people's emotions and facial expressions.

If I had autism, I would put the glasses on, and in real-time, the system would read the expressions of other people I was interacting with, and it would give me real-time, helpful cues. It would say, "Oh, you know what? You've been monologuing for the last 15 minutes. You may want to consider asking a question or taking a breath, because, you know, I think they're really bored."

**FAKE:** Rana's technology drew extensive interest from companies.

**EL KALIOUBY:** These Fortune 500 companies would say, "Have you thought of applying this in automotive, or for product testing, or content testing, or banking, or retail?" And I kept a log.

**FAKE:** When Rana's log reached 20 possible use cases, she and her mentor – the acclaimed professor Rosalind Picard – decided to launch Affectiva. Their goal was to make computers more human. And their sense of purpose really speaks to me. But technology is often built by people with good intentions. And then is appropriated by people with evil intentions. Rana's invention ran this risk. And she knew it.

**EL KALIOUBY:** When we spun out of MIT, my co-founder, Professor Rosalind Picard and I – we sat around her kitchen table in her house and we said, "Okay, we recognize that this technology can be used in many, many places. Where are we gonna draw the line?"

**FAKE:** Rana and Rosalind could imagine a lot of dark possibilities.

**EL KALIOUBY:** As an example, security and surveillance is a big area where we could, as a company, make a ton of money. But we decided that that was a space we were not gonna play in. And we got tested on it.

**FAKE:** Entrepreneurs are tested when the money gets tight. This is what happened to Rana and her team.

**EL KALIOUBY:** In 2011, we were literally a couple of months away from shutting down, we were running out of money. And we got approached by the venture arm of an intelligence agency and they wanted to put about \$40 million dollars in the company. But

they wanted to use the technology primarily for security surveillance and lie detection applications. And we had to take a stance.

**FAKE:** They turned the money down.

**FAKE:** A lot of entrepreneurs when faced with that, if they hadn't already discussed it, would not have had the power to have been able to say no.

**EL KALIOUBY:** I do feel like it's important that you be very clear around what your core values are. It's also our responsibility as leaders in this space to educate the public about all of the different use cases. Because what I like to say is: technology is neutral, right? Any technology in human history is neutral, it's how we decide to use it.

**FAKE:** And this seems a good place to pause so I can show you how this technology works. Today, Affectiva's technology reads human emotions through our face and our voice. With the voice, it's not about what we say. But how we say it.

**EL KALIOUBY:** How fast I'm speaking, how much energy is in my voice. Is it flat and monotonous or is it excited and kind of intense? Am I speaking fast or slow?

**FAKE:** And then, there's the face. All our expressions – smiles, smirks, grimaces – are controlled by the dozens of tiny muscles around our eyes, nose, and mouth. In the 1970s, scientists developed a coding system to categorize these expressions.

**EL KALIOUBY:** When you furrow your eyebrow, often an indicator of confusion or anger, that's action unit 4. A lip corner pull, which is pulling your lips outwards and upwards – which is what we would call a smile – is action unit 12.

**FAKE:** These “action units” combine in different ways to convey different emotions. And subtle differences can change the meaning. A smile is not always a smile.

**EL KALIOUBY:** If I pull both my lips upwards and outwards but I combine it with my eyes kinda shrink a little bit and they close a little bit that's usually a Duchenne smile. It's a true spontaneous smile of enjoyment. If I only smile with one half of my face, that's a smirk. And a smirk has very negative connotations.

**FAKE:** When Affectiva's technology scans an image of a face, it analyzes all these muscles in real-time to identify the “action units”. But what each facial action unit *means* – for different people in different cultures – is another story.

**EL KALIOUBY:** What is more tricky and more challenging is mapping that action unit 12 into an emotion word that we recognize as humans, and that we can relate to as a feeling. That science is, I'd say, less well understood. It's complex. Humans are complex.

**FAKE:** We wanted to help you picture this. So we tested Affectiva's software on a few iconic moments captured on video, and we're going to share the results.

Now to do this, we used Affectiva's iPhone app. It uses your phone's camera to detect faces and then overlays a digital readout on top of the face.

The digital readout tells you which emotion the face is showing. It tracks anger, disgust, fear, joy, sadness, and surprise – and the app gives you a percentage for the relative amount of each emotion. This percentage changes in real time with the facial expression.

We'll start with a piece of footage that delights me every time I see it – though it was admittedly not so delightful for the people involved. It's from a BBC newscast in 2017. An expert – via a live video feed from his home – is giving a somber assessment on the geopolitical situation in South North Korea. Then this happens.

**AUDIO:** “And what will it mean for the wider region? I think one of your children has just walked in.”

**FAKE:** A door in the background bursts open, and through it dances a young girl - she wants to see what daddy's doing. She's swiftly followed by a baby in a walker, and the mother just behind. As the expert tries to continue the interview, the mother dives through the door, scoops both toddlers into her arms and backs out of the room. It is glorious.

You don't need advanced AI to recognize the look of complete panic on mom's face.

What is interesting is the expert. There's one moment where he reaches behind and gently pushes his daughter back. It looks like he might be grimacing. But when you run that moment through Affectiva, it's not contempt that registers but 100% joy at his daughter's charming transgression – which makes me love the video even more.

Now for another live display that was a grab bag of heightened emotions. Miss Universe 2015. Host Steve Harvey on stage with the final two contestants: Miss Colombia and Miss Philippines. They stand center stage, holding each others hands. Steve opens the envelope and reads.

**STEVE HARVEY:** “Miss Universe 2015 is Colombia.”

**FAKE:** Miss Colombia's reaction is exactly what you would expect - and Affectiva registers it as 100% joy. She gets a sash, a crown, flowers. She waves a small Colombian flag, blows kisses. One hundred percent joy. But then Steve Harley sidles onto stage. You can tell something's up.

**HARVEY:** “Listen folks. Let me just take control of this. I have to apologize.”

**FAKE:** Focused on Steve's face, Affectiva registers a 46% spike in fear. A millisecond later it registers anger at around 20% as he waits for the crowd to quiet down. He looks down at his cue card. Disgust jumps to 80%. Then he makes the announcement.

**HARVEY:** "The first runner-up is Colombia. Miss Universe 2015 is Philippines."

**FAKE:** The camera cuts to Miss Philippines. The meaning dawns on her. Her fear turns to 100% joy. As for Steve and Miss Colombia – well, I can't tell you, because by this time Steve made a rapid exit.

On *Should This Exist?*, we're creating a new kind of conversation – a conversation between the entrepreneur and the world.

Welcome to the *Should This Exist?* workshop. Here, Rana and I will respond to ideas I've gathered from super smart, creative experts. We asked each of them to throw unexpected things at us – both possibilities and pitfalls.

We have a stacked roster of experts today, and we'll start with Joy Buolamwini, a researcher at the MIT Media Lab focusing on facial recognition software and founder of the Algorithmic Justice League, who also happens to be a poet.

In 2019, we can't talk about facial recognition without talking about the potential for algorithmic bias. And this is Joy's biggest worry about Affectiva's AI tools.

Joy has spent years testing the accuracy of systems built by companies like Microsoft and Amazon. She'll feed the system photos – mainly of faces – and test the algorithm. The results are often shocking.

**JOY BUOLAMWINI:** We had Microsoft captioning Michelle Obama as a boy smiling at a camera in one of her photos.

**FAKE:** Wow.

**BUOLAMWINI:** We had Serena Williams being labeled male. Amazon Rekognition, which is being sold to law enforcement right now – we ran Oprah Winfrey's face through their system. It's guessing her to be male. And I'm like, "Wait, wait, wait, wait, if we can't get Oprah, correct – are we sure we want to be selling this to law enforcement?" And earlier this year, announced that Amazon Rekognition it is working with the FBI.

**FAKE:** Oh gosh.

**FAKE:** Along with this quantitative data, Joy also collects personal stories from people whose everyday lives are impacted by facial recognition gone wrong. The stories are often submitted anonymously, or on behalf "of a friend".

**BUOLAMWINI:** One submission we got said “I have a friend who works at a large tech company, and she's limited to only booking conference rooms that have a face recognition system that works with her. So this limits her workplace productivity.

**FAKE:** Unbelievable.

**BUOLAMWINI:** Right?

**FAKE:** And so this is something that presumably a large tech company if it's got like that kind of tech.

**BUOLAMWINI:** It's a large tech company you've probably heard of.

**FAKE:** The idea of an African-American woman in 2019 running around to find a conference room that can recognize her face is just... appalling.

**BUOLAMWINI:** The company kind of outed themselves by saying, “Oh we're aware of this and we're trying to make changes.”

**FAKE:** Now, when you take facial recognition – which already, typically struggles to identify people who aren't white – and add a layer of interpretation on top of that? This is what worries Joy about Affectiva.

**BUOLAMWINI:** Can you measure an internal state from somebody's external expressions? And this is contested, right? So this is where a company like Affectiva can start facing criticisms.

**FAKE:** So, yes, you know, we've seen this everywhere. How do you as a facial recognition AI company guard against this happening?

**EL KALIOUBY:** Absolutely. So first of all, I'll start by saying I'm a huge fan of Joy and her work, and I think we need more people like her to advocate against algorithmic and data bias. The way to think about this is that it's a combination of the data you use for training these algorithms and then the actual algorithm. And I would argue the team designing for these algorithms as well.

If you feed the algorithm faces or emotion expressions of middle-aged white men, that's what the algorithm's going to learn. And then if you then deploy it on darker skin colored people or females in Asia – it hasn't really seen examples of that.

The more diverse the people how are designing these systems are, the more they can say, "You know, I noticed that there aren't enough data of people with my skin color, darker skin color. Can we make sure we include that?" Or, "I have a beard and I notice that we don't have any people in this data set that have beards," and the list goes on and on and on.

**FAKE:** Have you seen instances of this bias in any of the Affectiva research that you've done so far, in some of your prototypes or anything that you've worked to correct?

**EL KALIOUBY:** Yeah, when we first started – over seven years ago – the algorithm, had a small repertoire of emotions that it could detect. And I remember we had a number of really key clients in China and they reached out to us and they said, “Okay, this algorithm is rubbish, it doesn't really work on Chinese people.”

And we were like, “Hmm. I wonder what's going on there.” And we kind of dug into the data and we noticed that a lot of Asians have this very subtle smile. It's kind of a very subtle lip corner pull. It's not a happiness smile. It's a politeness or social smile. And we didn't have examples of that social smile in our training data set, and so to fix it we went back and we added samples of Asian people doing this kind of very subtle smile to our data set – and that fixed the problem.

**FAKE:** Some worry rightly about the accuracy of Affectiva's algorithms. But Sam Altman sees a future where the algorithm becomes frighteningly accurate. Sam is the chairman of Y Combinator and also co-founder of Open AI, a non-profit research company whose mission is to advance AI in a way that benefits humanity. As a lifelong science fiction fan, I was intrigued by Sam's first thought.

**SAM ALTMAN:** One thing I fundamentally believe is that people have a right to their own thoughts. And another thing I believe is that this will eventually be better than humans at reading emotion. You know, there'll be these micro expressions that flash across the face that even humans don't quite pick up on – like as good as we are, as well attuned as we are – very tiny, very fast, very subtle facial changes that are below my detection threshold on you sitting this far across the room.

What happens when you can get in trouble for “wrong think” because I said something in a flash of anger that I missed myself, but the AI was right about what flashed across your face? And then you recovered and did the right thing.

**FAKE:** Right. And then you're into like “precrime” and Minority Report stuff.

**ALTMAN:** Totally.

**FAKE:** Here's what Rana had to say about that.

**EL KALIOUBY:** We're not in the business of reading your thoughts. So yes, I can detect that you just nodded your head or you smiled. I have no idea what's going on in your head – and I think that should stay private.

**FAKE:** But I think what Sam is saying is that it's very difficult to mask your emotions. As you said earlier, some ninety-four percent of all of our communication is nonverbal. And so much of that kind of communication is hard to disguise. I think what Sam is saying is that you should be defended against that and somehow you need to have the privacy of your own thoughts.

**EL KALIOUBY:** Right. And I would hope that we are in a world where you are motivated to share that 93% instead of you feel obligated to mask. If we craft this technology and the applications in a way where you're like, "Oh my God, I'm just going to have a poker face all the time," then I think we failed. We have failed as AI designers and innovators.

But on the other hand if we craft in a way where there's value in you sharing these amazing signals, then I think that's a really exciting world.

**FAKE:** I do think that there are contexts in which you're more likely to see the benefit than others – if it's a safety application in a car, that's very different than if it's a facial recognition on your refrigerator providing you with more ice cream. Is it advertising, is it kind of pushing products on you? And that's like the entrepreneur's choice. Which of those directions do you pursue? And sometimes the less ethical one is the more lucrative one.

**EL KALIOUBY:** So I'm an optimist. I do think part of what you're saying is right. I do think that there is, for our technology, there's probably a clear path to being very profitable in security and surveillance. But that's a place where we chose not to play.

**FAKE:** Affectiva was designed to mediate human relationships in the digital age. So we wanted to hear from Esther Perel, the renowned relationship expert. Esther is the author of *State of Affairs*, and host of the hit podcast *Where Should We Begin?* She's known for integrating technology into her work with patients. But Affectiva raised some immediate flags for Esther.

**PEREL:** So many people would rather send messages than talk. They would rather text than talk. Why will I use a technology that will help you understand the emotions of the person who's writing to you, instead of actually walk five meters and go talk to your neighbor? And what you need to learn is to re-engage with people and deal with the discomfort, which will be actually lesser.

The more we insulate ourselves from the physical, messy, interactive, iterative processes of relating to others, the more we will need to develop technologies that help us deal with the discomforts. But we have created those very discomforts or amplified those very discomforts by minimizing friction, by minimizing all the situations in which people used to deal with others.

**FAKE:** That's such a good perspective. It's super interesting.

**EL KALIOUBY:** To Esther's point, we're not really practicing using our nonverbal communication skills. We're not really practicing making eye contact. We're not practicing facing conflict. And as a result, we're not getting better at it.

I thought Esther was going to talk about relationships, like romantic relationships, because that's an area where I think there's a lot of potential use for this technology.

**FAKE:** What could it solve for?

**EL KALIOUBY:** I mean, I think it could be the bridge – if you're sending off really angry vibes, it could flag that. It could say: “You were actually yelling today. You think you weren't, but you know what? You were really yelling.” Or: “You weren't really paying attention during this conversation. You were half present.” So I think it could provide a third-party objective arbitrator for some of these conflicts.

**FAKE:** “You were kind of absent...”

**EL KALIOUBY:** Right. Or condescending or disrespectful.

**FAKE:** Yeah. And what if it's an app that both partners have? I mean, I'm just trying to like imagine this, kind of great...

**FAKE:** Let's look at another utopian scenario. As the director of the MIT Media Lab, Joi Ito is constantly looking at emerging technologies, and predicting their consequences. Joi was quick to spot a range of possible outcomes for Affectiva.

**JOI ITO:** The positive side would be: I'd love for – with consent – to have children and teachers connect so that you can improve learning outcomes. And if it's a mother with a child, you can understand how children are feeling – and maybe alert caregivers for potentially dangerous situations.

I think there are so many times when somebody wishes they could express their emotions and can't, and I think that a lot of this comes back to the genesis story of the technology itself.

**FAKE:** This resonated with me. As a parent, I would have loved to use Affectiva to understand what my newborn baby was crying about. Rana was also excited about this potential use case.

**EL KALIOUBY:** So there's a lot of exciting use cases in what Joi said. You could imagine with technology like ours and a ton of data, you could actually develop a code book for different crying patterns: Is the baby hungry? Is the baby upset? Does the baby need a diaper change? So that's one scenario.

I'm especially excited about the applications in education. You could quantify the level of attention: Is the student fidgeting? Or are they confused or bored? And then you have

the learning system personalized and adapt in real time, just like an awesome teacher would.

**FAKE:** I get really excited about scenarios – like online education – where Rana’s technology makes devices more human, more empathic. But as I said earlier, these bright ideas cast a dark shadow. You can imagine a future where systems that know how you feel can more easily manipulate you.

This idea came up in different ways with both Esther, the couples therapist, and Joi, the director of the MIT Media Lab. Here’s what Joi had to say:

**ITO:** You want to go to the gym. You don't want to want Oreos at midnight. But you can't actually control that. But advertising, well-placed and talking to your subconscious and nudging you using all kinds of feedback systems – can start to make you want things that the advertiser wants you to want – but that you don't want to want. The old-fashioned word is “brainwashing,” right?

**FAKE:** And here’s Esther, who offered a similar sentiment.

**PEREL:** In fiction – in science fiction – the thought of entering another person's mind to be able to manipulate it, to be able to make it think that it is making its own decisions when in fact they have been engineered – it comes back to the fundamental question of free will or free choice. Do we have any?

Because we are homosapiens we think that we are making our own decisions, but in fact they have been engineered. Now, they can be engineered by tradition. They can be engineered by social pressure. They can be engineered by public shaming. They can be engineered by technologies.

**EL KALIOUBY:** I agree with Esther. I can see a version of the world where this technology could be used to manipulate or exploit your emotional state to persuade you one way or another to buy a product, vote for somebody. But I also want to argue that we're already in this world. Our data is already being used to target ads to us, right?

**FAKE:** I always find it completely shocking whenever I arrive at the airport in Las Vegas: There is disorienting carpet. There are flashing lights. There are no signs of what time of day it is. It is completely dark. And all of that is designed to put you into a state of disorientation and vulnerability and susceptibility to the desire to pull the lever on the “One-Armed Bandit” one more time, right? I mean, there are environments like that.

**EL KALIOUBY:** We're already in that world.

**FAKE:** But I think that what Esther and Joi are afraid of is that this could take it so much further.

**EL KALIOUBY:** I do worry about that. I actually particularly worry about that in the kind of political advertising context. So you could imagine how this could be abused in the hands of the wrong leaders.

**FAKE:** What would the dystopian scenario be?

**EL KALIOUBY:** I mean, I imagine in the hands of autocratic leaders, if you see a political message and you do a smirk – often an indicator of skepticism or doubt, so now you're doubting this leader – I imagine how that could be used against you.

**FAKE:** And what if it gets into our homes?

**EL KALIOUBY:** Again, I think if it gets into our homes, there is huge potential for good, like it could be your friend or companion that helps you be more active and sleep better and eat healthier.

I think there's a lot of potential where this thing knows you so well, it can help you become a better version of you. But that same data in the wrong hands could be could be manipulated to exploit you.

**FAKE:** If we're going to guide Rana's technology – and AI more broadly – to its best possible future, we need to involve more than one company – and in fact more than one country in the conversation. Greg Brockman has some ideas on this. He's the co-founder, with Sam Altman, of OpenAI. Greg wonders whether companies like Affectiva can play a role in finding common ground.

**GREG BROCKMAN:** One thing I think is interesting is the difference of just popular opinion in different standards across different countries. In China, facial recognition is pretty prevalent right? It's really ubiquitous and used for a number of different really interesting applications. And here, I think, that there's much more push back and more reticence to deploy it.

Think about seat belts. Everyone has a seatbelt. Seat belts are just a good thing. And it's a good thing for everyone to coordinate on seatbelts and to make sure that everyone has the latest seatbelt technology, that everyone's thinking about how they can make the car safer and better. And the same is true in AI: that safety is something that I think can be the starting point for international coordination.

**FAKE:** There's a lot in what he said, but one of them is that there's different standards across different countries. And one of the things that we had talked about was how much more accepted this technology is in China and less acceptable here. But that ultimately we all agree on certain basic principles, of what it should and shouldn't do.

**EL KALIOUBY:** I do think there are cultural differences in how people approach AI in general, and in particular, our kind of our subcategory of AI, understanding all things human.

And we find, in our experience, that that is not necessarily shared by other countries like China – which actually has real implications on the development of the technology – because it's all data-driven, the organization that has access to the most data ultimately potentially wins, right?

**FAKE:** Who do you think right now has the most data? Which company, or perhaps which government?

**EL KALIOUBY:** I definitely think the Chinese government has access to the most face video data. So China is taking a different approach, which on the surface puts them ahead potentially – because they have access to more data and more data across different devices and scenarios and applications, which makes for better algorithms. But these core values don't don't match our core values.

**FAKE:** So in some ways it sounds like you're you're disagreeing with Greg and Sam a bit because they think that when you look at it from the 35,000 foot-level we agree, but maybe we don't.

**EL KALIOUBY:** Yeah, I actually love the seat belt example that Greg gave, because we're seeing that in the automotive space, there is cross-cultural agreement that we need some sensors inside the vehicle – interestingly China is really pushing for regulation for this kind of passive safety monitoring of drivers and passengers – so, I think there's agreement there. I just think broadly there may not be agreement on the approach for how to get there.

**FAKE:** There's clearly a lack of agreement, between countries and companies on acceptable uses of A.I. And this points to a great dilemma, which Joi articulates.

**ITO:** I think the tricky thing about technology is that if you don't do it, somebody else can in some ways. And so even if Affectiva doesn't do it, once people know that it can be done, that creates a business opportunity for somebody to do something that you wouldn't do.

And I think we saw this even with nuclear energy and the atom bomb. I think a lot of the scientists who are involved in the early days started to have misgivings, but once it was out, they didn't have that much power to uninvent the thing they had invented, right?

**FAKE:** “The power to un-invent the thing they invented.” I love the words Joi put to that idea – it's one I've thought about a lot. And so has Rana.

**EL KALIOUBY:** I like to think that the technology we're building has a lot of applications that are truly transformative, that can really really help people – but there's definitely potential for abuse in all sorts of ways.

And for us as a company, we've chosen to be very clear about our licensing terms. So if you're going to license our technology, you cannot use it for use cases where people are not consenting. And then I just like to be a public advocate for how this technology could be used for good or for evil, and I feel like that's our social responsibility to educate the public. But you're right: Once you invent a technology, it's really hard to uninvent it. Joi's right.

**FAKE:** Rana's technology was fascinating to explore because of how extreme its use cases are. For example, if administered in classrooms, it can be transformative in students' lives.

On the other hand, we can create a real life version of 1984 where your private thoughts can get you in trouble. I'm still working through the implications myself.

Rana puts forth a compelling case for her tech, but we want to know what you think. Tweet us using #ShouldThisExist; and rate, review and subscribe on Apple podcasts or wherever you listen.