



# **BUILDING ETHICAL, EMPATHIC, AND INCLUSIVE AI**

**Rana el Kaliouby, PhD**

**It is commonplace** on social media and in politics, entertainment, and popular culture to see callous, hateful language and actions that even a couple of years ago would have been considered shocking, disgraceful, or disqualifying. As a newly minted American citizen born in Egypt and a Muslim woman who immigrated to the United States at a time when political leaders were calling for Muslim bans and border walls to keep immigrants out, I am particularly aware of the insensitive, at times vicious, voices in the cyber world. But, in truth, everyone is fair game.

It is not off-limits to troll survivors of gun violence, like the students of Parkland, who advocate for saner gun laws; to shame victims of sexual abuse; to post racist, anti-Semitic, sexist, homophobic, anti-immigrant rants; or to ridicule people whose only sin is that they disagree with you. This is happening in our communities, our workplaces, and even on our college campuses. Today, such behaviors are dismissed with a shrug. They can even get you tens of millions of followers and a prime-time spot on a cable TV network—or send you to the White House.

It is all representative of a problem endemic to our society. Some social scientists call it an “empathy crisis.” It is the inability to put yourself in another’s shoes and feel compassion, sympathy, and kinship for another human being. This stunning lack of concern for our fellow citizens permeates and festers in the cyber world, especially on social media, and is spilling over into the real world.

We are in increasingly dangerous territory—we are at risk of undermining the very traits that make us human in the first place.

More than two decades ago, journalist Daniel Goleman wrote about the importance of empathy in his bestselling book *Emotional Intelligence*. When he argued that genuine intelligence is a mixture of IQ and what we've come to call EQ, or emotional intelligence, he changed our thinking about what makes someone truly intelligent. EQ is the ability to understand and control our emotions and read and respond appropriately to the emotional states of others. EQ, rather than IQ, is the determining factor predicting success in business, personal relationships, and even our health outcomes.

Obviously, you can't experience emotional intelligence without feeling, without emotion. But when we are in cyberspace, "feelings" don't come into play because our computers can't see or sense them: When we enter the virtual world, we leave our EQ behind.

Inadvertently, we have plunged ourselves headfirst into a world that neither recognizes emotion nor allows us to express emotion to one another, a world that short-circuits an essential dimension of human intelligence. And today, we are suffering the consequences of our emotion-blind interactions.

**We are in increasingly dangerous territory—we are at risk of undermining the very traits that make us human in the first place.**

Computers are “smart” in that they were designed with an abundance of cognitive intelligence, or IQ. But they are totally lacking in EQ. Traditional computers are emotion blind: They don’t recognize or respond to emotion at all. Around twenty years ago, a handful of computer scientists—I was one of them—recognized that as computers became more deeply embedded in our lives, we would need them to have more than computational smarts; we’d need them to have people smarts. Without this, we run the risk that our dependence on our “smart” technology will siphon off the very intelligence and capabilities that distinguish human beings from our machines. If we continue on this path of emotion blind technology, we run the risk of losing our social skills in the real world. We will forget how to be compassionate and empathetic to one another.

## THE NEW REALITY

I am cofounder and CEO of a Boston-based artificial intelligence company, a pioneer in Emotion AI, a branch of computer science dedicated to bringing emotional intelligence to the digital world. My goal is not to build emotive computers, but to enable human beings to retain our humanity when we are in the cyber world—to humanize technology before it dehumanizes us.

In striving to become the “expert” I needed to be in human emotion in order to teach machines about emotion, I found myself turning the spotlight on my own emotional life. This was an even more daunting process than writing code for computers; it forced me to confront my own reticence to share my innermost feelings, indeed, my own reluctance to recognize and act on my own feelings. Ultimately, decoding myself—learning to express my own emotions and act on them—was the biggest challenge of all. Expert as I have become on the subject, I feel that I am very much a work in progress myself.

To me, my work and my personal story are inseparable; each flows into the other. I am a rarity in the tech world: a woman in charge—a brown-skinned computer scientist at that—in a field that is still very male and very white. I was raised in the Middle East in a male-dominated culture that is still figuring out the role of women in a world that is changing with breathtaking speed. In both these cultures—tech and the Muslim Middle East—women have been excluded or restricted from positions of power. I’ve had to maneuver both cultures to achieve what I have.

Bridging the divide between people is how we gain empathy, and that is how we build strong, emotionally intact people and a strong, emotionally intact world. That is the core of what I do, whether in the real world or the cyber one. I am also passionate that people understand what AI is, and how it’s going to impact their lives. The world, your world, is about to change. Given how AI is becoming ubiquitous, and the potential impact it has on all our lives, it’s critical that we as a society take an active role in how this AI is designed, developed, and deployed.

By this year, according to various industry reports, there will be between four and six connected devices for every human being on the planet. And our computers will become even more embedded in our lives. This is the new reality.

I am not suggesting that face-to-face relationships are not still important—just the opposite. Yes, it’s unacceptable to sit at dinner texting instead of engaging in real conversation with the people around you, but the reality is that much of your interpersonal interactions today are conducted in the cyber world, and that’s not going to change. (I’m a pragmatist.)

So the solution is not to turn back the clock, shut down our devices, and go back to life as it was before our computers. We depend too much on technology to give it up. And doing so would be a terrible mistake. We need our technology more than ever. But we need to make it smarter, better, and more humane. And fortunately, we now have the tools to do this.

An Emotion AI world is a human-centric world, one where our technology has our backs and helps us become healthier, happier, and more empathetic individuals: technologies such as Google Glass, equipped with “emotion decoders” that help autistic kids better interact socially with others; semi-autonomous cars that assume control of the wheel when we’re too angry, distracted, or tired to drive safely. Preventing millions of accidents each year; emotion-aware devices (from smartwatches and smartphones to smart refrigerators) able to detect mental and physical ailments years before their onset; empathetic virtual assistants that can track your mood and offer timely guidance and support; human resources emotion analytic tools that can enable an HR recruiter to match the right person more precisely to the right job or team and eliminate much of the unconscious bias that occurs in hiring; and an intelligent learning system that can detect the level of engagement of a student and tailor its approach accordingly.

**We need our technology more than ever.  
But we need to make it smarter, better, and  
more humane.**

The potential for Emotion AI is breathtaking, but I am not naive: When you have computers with the capacity to recognize and record the emotional states of users, of course privacy is of major concern. Emotion AI should—must—be employed only with the full knowledge and consent of the user, who should be able to opt out at any time. Emotion AI will know a lot about us: our emotional states, moods, and interactions. In the wrong hands, this information can be very damaging. That is why it is so essential for the public to be aware of what this technology is, how and where data is being collected, and to have a say in how it is being used.

It is also imperative that AI technology be developed with all human beings in mind, that it be inclusive. Our software must reflect the real world, not just the world of an elite. Much of my research has been devoted to obtaining data from a diverse population, representing all ages, genders, ethnicities, and geographical areas. If our AI fails to do this, we will be creating a new form of discrimination that will be very hard to undo, and as such technology propels us forward, we run the risk of leaving behind whole sectors of our population, and that would be disastrous for us all.

## **HUMAN BEFORE ARTIFICIAL**

As AI becomes mainstream and ubiquitous, AI systems that are designed to engage with humans will have a lot of data on you—personal data about who you are, your preferences, your actions and your quirks. We live in a society where data is being collected on all of us, all the time. Sometimes it's obvious; sometimes not. Sometimes it's for our benefit; sometimes not. As the CEO of an AI company, I understand that having access to so much data comes with a great deal of responsibility: How do we ensure the ethical development and deployment of AI?

My team and I offer a positive vision for AI, but we're not naive. We understand that the potential for abuse, especially the careless management of data, is very real and it is happening right now, but companies and governments.

Privacy advocates have long voiced concerns over how data is collected by Big Tech, and how it can be sold or hacked, or can inadvertently end up in places where it doesn't belong. Technology companies have a long history of acting first and asking for forgiveness later. Lately, the public is not so forgiving. But let's face it, every time we search the Internet, shop, or download a book or video, we are being watched. Like it or not, a lot of people whom we don't know know an awful lot about us and our preferences. Stalking you online while you surf your favorite stores to learn your brand preferences (without your permission) may be an infringement of your privacy, but it pales in comparison to how totalitarian regimes have weaponized AI against entire groups of people. Totalitarian regimes ignore the public because they can. But in a global marketplace, the "public" isn't just your own citizens. You have to consider the citizens of the world. We can build a better AI on our strengths and ethical standards, one that empowers individuals. At the same time, we can reform Big Tech and build a more humanistic AI for the future.

AI is not evil. The technology itself is neutral, but it may be used by people with nefarious purposes. We (software developers) have a responsibility to be highly selective about whom we allow to use our technology, and how it will be used. This is where consumers can flex their muscles: Do you want to buy a product from a company that allows its AI to be used to spy on ethnic minorities, as it is used in China? Consumers have power that they think to control abuse in this sector. They just haven't had the tools to take action yet.



Big Tech companies are beginning to recognize that if they don't take action to prevent abuses, local and federal governments will. Certainly, some regulation is required. But these issues are so complex, and the technology moving so quickly, that the industry itself has to step in and take a proactive role in designing a strategy that doesn't inhibit progress, but also doesn't achieve it at the cost of privacy. It is one reason my company, Affectiva, is part of a Partnership on AI to Benefit People and Society (PAI), and tech industry consortium established to "study and formulate best practices on AI technologies, to advance the public's understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influence on people and society." Its eighty-plus members from more than thirteen countries include tech giants as well as other voices, like the American Civil Liberties Union, Amnesty International, and the Hastings Center, a bioethics research institute. It's why we started our annual Emotion AI Summit with the goal of building an ecosystem of ethicists, academics, and AI innovators and practitioners across industries to take action.

As a society, we are just starting to have a conversation about the role of AI in our lives, and how to use it ethically and fairly for the betterment of humanity. We can't let the presence of some bad players exploiting this technology thwart the ability of the good players to create tools and services that are helpful to society. We have to establish standards, and those who cross the line, either directly or indirectly (by licensing technologies to bad players), need to know that the tables have turned: AI developers, consumers, and different levels of government are now tracking them and will penalize companies that don't conform to basic human standards.

**After all, shouldn't we put the human before artificial? 🤖**



# Info



Ready to dig deeper into this idea?  
Buy a copy of [Girl Decoded](#).

Want copies for your organization or for an event? We can help:  
[customerservice@porchlightbooks.com](mailto:customerservice@porchlightbooks.com) 800-236-7323

## ABOUT THE AUTHOR

Rana el Kaliouby is a pioneer in artificial emotional intelligence (Emotion AI), as well as the co-founder and CEO of Affectiva. She grew up in Cairo, Egypt, among a family of educators. After earning an undergraduate and master's degree in computer science at the American University in Cairo, she attended Cambridge University, where she earned her Ph.D. Afterwards, she joined the MIT Media Lab as a research scientist, where she spearheaded the applications of emotion recognition technology in a variety of fields, including mental health and autism. She left MIT to co-found Affectiva, the company credited with defining the field of Emotion AI. Her company now works with more than a quarter of the companies in the Fortune Global 500. An acclaimed TED speaker, Rana was named by *Forbes* to their list of America's Top 50 Women in Tech, and *Fortune* included her in their list of 40 Under 40. In 2018 she was the cohost of a PBS Nova series on AI.



## Porchlight

Curated and edited by the people of Porchlight, ChangeThis is a vehicle for big ideas to spread. Keep up with the latest book releases and ideas at [porchlightbooks.com](http://porchlightbooks.com).

This document was created on April 22, 2020 and is based on the best information available at that time.

The copyright of this work belongs to the author, who is solely responsible for the content. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License. To view a copy of this license, visit Creative Commons. Cover image from Adobe Stock.

### SHARE THIS

Pass along a copy of this manifesto to others.



### SUBSCRIBE

Sign up for e-news to learn when our latest manifestos are available.

