

'Human Beings Are Soon Going to Be Eclipsed'

July 13, 2023



Recently I stumbled across an essay by Douglas Hofstadter that made me happy. Hofstadter is an eminent cognitive scientist and the author of books like "Gödel, Escher, Bach" and "I Am a Strange Loop." The essay that pleased me so much, called "The Shallowness of Google Translate," was published in The Atlantic in January of 2018.

Back then, Hofstadter argued that A.I. translation tools might be really good at some pedestrian tasks, but they weren't close to replicating the creative and subtle abilities of a human translator. "It's all about ultrarapid processing of pieces of text, not about thinking or imagining or remembering or understanding. It doesn't even know that words *stand* for things," he wrote.

The article made me happy because here was a scientist I greatly admire arguing for a point of view I've been coming to myself. Over the past few months, I've become an A.I. limitationist. That is, I believe that while A.I. will be [an amazing tool](#) for, say, tutoring children all around the world, or summarizing meetings, it is [no match for human intelligence](#). It doesn't possess understanding, self-awareness, concepts, emotions, desires, a body or biology. It's bad at [causal thinking](#). It doesn't possess the nonverbal, tacit knowledge that humans take for granted. It's not sentient. It does many things way faster than us, but it lacks [the depth of a human mind](#).

I take this to be good news. If A.I. is limited in these ways, then the A.I. revolution will turn out to be akin to the many other information revolutions that humans have produced. This technology will be used in a lot of great ways, and some terrible ways, but it won't replace us, it won't cause the massive social disruption the hypesters warn about, and it's not going to wake up one day wanting to conquer the world.

Hofstadter's 2018 essay suggested that he's a limitationist too, and reinforced my sense that this view is right.

So I was startled this month to see the following headline in one of the A.I. newsletters I subscribe to: "Douglas Hofstadter Changes His Mind on Deep Learning & A.I. Risk." I followed the link to a podcast and heard Hofstadter [say](#): "It's a very traumatic experience when some of your most core beliefs about the world start collapsing. And especially when you think that human beings are soon going to be eclipsed."

Apparently, in the five years since 2018, ChatGPT and its peers have radically altered Hofstadter's thinking. He continues: It "just renders humanity a very small phenomenon compared to something else that is far more intelligent and will become incomprehensible to us, as incomprehensible to us as we are to cockroaches."

I called Hofstadter to ask him what was going on. He shared his genuine alarm about humanity's future. He said that ChatGPT was "jumping through hoops I would never have imagined it could. It's just scaring the daylights out of me." He added: "Almost every moment of every day, I'm jittery. I find myself lucky if I can be distracted by something — reading or writing or drawing or talking with friends. But it's very hard for me to find any peace."

Hofstadter has long argued that intelligence is the ability to look at a complex situation and find its essence. "Putting your finger on the essence of a situation means ignoring vast amounts about the situation and summarizing the essence in a terse way," he said.

Humans mostly do this through analogy. If you tell me that you didn't read my column, and I tell you I don't care because I didn't want you to read it anyway, you're going to think, "That guy is just bloated with sour grapes." You have this category in your head, "sour grapes." You're comparing my behavior with all the other behaviors you've witnessed. I match the sour grapes category. You've derived an essence to explain my emotional state.

Two years ago, Hofstadter says, A.I. could not reliably perform this kind of thinking. But now it is performing this kind of thinking all the time. And if it

can perform these tasks in ways that make sense, Hofstadter says, then how can we say it lacks understanding, or that it's not thinking?

And if A.I. can do all this kind of thinking, Hofstadter concludes, then it is developing consciousness. He has long argued that consciousness comes in degrees and that if there's thinking, there's consciousness. A bee has one level of consciousness, a dog a higher level, an infant a higher level, and an adult a higher level still. "We're approaching the stage when we're going to have a hard time saying that this machine is totally unconscious. We're going to have to grant it some degree of consciousness, some degree of aliveness," he says.

Normally, when tech executives tell me A.I. will soon achieve general, human level intelligence, I silently think to myself: "This person may know tech, but he doesn't really know human intelligence. He doesn't understand how complex, vast and deep the human mind really is."

But Hofstadter *does* understand the human mind — as well as anybody. He's a humanist down to his bones, with a reverence for the mystery of human consciousness, who has written movingly about love and the deep interpenetration of souls. So his words carry weight. They shook me.

But so far he has not fully converted me. I still see these things as inanimate tools. On our call I tried to briefly counter Hofstadter by arguing that the bots are not really thinking; they're just piggybacking on human thought. Starting as babies, we humans begin to build models of the world, and those models are informed by hard experiences and joyful experiences, emotional loss and delight, moral triumphs and moral failures — the mess of human life. A lot of the ensuing wisdom is stored deep in the unconscious recesses of our minds, but some of it is turned into language.

A.I. is capable of synthesizing these linguistic expressions, which humans have put on the internet and, thus, into its training base. But, I'd still argue, the machine is not having anything like a human learning experience. It's

playing on the surface with language, but the emotion-drenched process of learning from actual experience and the hard-earned accumulation of what we call wisdom are absent.

In a piece for *The New Yorker*, the computer scientist Jaron Lanier argued that A.I. is best thought of as “an innovative form of social collaboration.” It mashes up the linguistic expressions of human minds in ways that are structured enough to be useful, but it is not, Lanier argues, “the invention of a new mind.”

I think I still believe this limitationist view. But I confess I believe it a lot less fervently than I did last week. Hofstadter is essentially asking, If A.I. cogently solves intellectual problems, then who are you to say it’s not thinking? Maybe it’s more than just a mash-up of human expressions. Maybe it’s synthesizing human thought in ways that are genuinely creative, that are genuinely producing new categories and new thoughts. Perhaps the kind of thinking done by a disembodied machine that mostly encounters the world through language is radically different from the kind of thinking done by an embodied human mind, contained in a person who moves about in the actual world, but it is an intelligence of some kind, operating in some ways vastly faster and superior to our own. Besides, Hofstadter points out, these artificial brains are not constrained by the factors that limit human brains — like having to fit inside a skull. And, he emphasizes, they are improving at an astounding rate, while human intelligence isn’t.

It’s hard to dismiss that argument.

I don’t know about you, but this is what life has been like for me since ChatGPT 3 was released. I find myself surrounded by radical uncertainty — uncertainty not only about where humanity is going but about what being human is. As soon as I begin to think I’m beginning to understand what’s happening, something surprising happens — the machines perform a new task, an authority figure changes his or her mind.

Beset by unknowns, I get defensive and assertive. I find myself clinging to the deepest core of my being — the vast, mostly hidden realm of the mind from which emotions emerge, from which inspiration flows, from which our desires pulse — the subjective part of the human spirit that makes each of us ineluctably who we are. I want to build a wall around this sacred region and say: “This is essence of being human. It is never going to be replicated by machine.”

But then some technologist whispers: “Nope, it’s just neural nets all the way down. There’s nothing special in there. There’s nothing about you that can’t be surpassed.”

Some of the technologists seem oddly sanguine as they talk this way. At least Hofstadter is enough of a humanist to be horrified.

David Brooks has been a columnist with The Times since 2003. He is the author of “The Road to Character” and, most recently, “The Second Mountain.” [@nytdavidbrooks](#)

A version of this article appears in print on July 14, 2023, Section A, Page 22 of the New York edition with the headline: ‘Human Beings Are Soon Going to Be Eclipsed’. [Order Reprints](#) | [Today’s Paper](#) | [Subscribe](#)