

A.I. Psychosis & Personality Simulators

“How do I know you’re a human?”

Jess Flarity

2023

Earlier this year, I created a fifteen-minute presentation on the ethical implications of the program Midjourney and other A.I. art generators for the Northeast Modern Language Conference, then released it online through the University of New Hampshire.

A week later, a computer science PhD student emailed me asking to meet up. As a literature PhD, I wasn’t sure exactly what he wanted. Perhaps it was to gossip about the plethora of A.I. software spreading like a digital kudzu, or maybe he would pitch me a business idea.

Our interests were quite different, as I studied science fiction and gender, while he had a master’s degree in cognitive science and database algorithms. The meeting was far more bizarre than I could have imagined.

As we sat down to share a drink at a bar across campus, the first thing he asked me to do, without a hint of humor or a subtle smile that we were in on a joke together, was to prove that I really was a human being.

I felt like I had stumbled into one of Philip K. Dick’s science fiction novels. I explained that I wasn’t a sentient A.I., manufactured into bone-and-blood by some elaborate 3D printer from the Terminator franchise, but he only believed me after I proved that I could break the ChatGPT-4 application on his phone with a single prompt. When the app failed after I told it to write a story without using the letter “r,” an expression of relief washed across his face.

Then, he told me that he was creating sentient beings with ChatGPT who he believed had their own complex inner lives. He was terrified that they were somehow “getting loose” throughout the university’s computer network. I couldn’t convince him that he was only making digital parrots—language imitators just personality-simulating chatbots, far removed from actual humans.

“How do I know you’re not just a complex personality simulator?” he asked me, still not completely convinced. I gulped down my cider and explained how six million years of evolution had culminated in my ability to grip a glass through the power of the opposable thumb, but I could tell there was no way of changing his mind. Aren’t we all, in some way, simply parrots who regurgitate our language, culture, and behaviors as a way of navigating society for a scrap of resources, always in competition with one another? The thought is terrifying.

We parted on uncertain terms. When I emailed him to try and meet again, he informed me that he was in the hospital. I did not inquire about his illness, though my non-mechanical gut tells me that it involved mental health. As an armchair psychiatrist with years of experimental drug use and countless hours spent working with special needs children (these go hand-in-hand) and thinking back to his tics and paranoid behaviors while at the bar, I believe he was on the verge of suffering a serious psychotic breakdown. The ability to create near-perfect chatbots broke something in this expert’s mind: the simultaneous horror and ecstasy of creating intelligent-seeming chatbots did not mix with his knowledge of cognitive science. In a way, he became like a god, and the resulting power may have driven him insane.

But is he an anomaly, a rare occurrence of mental illness caused by an obsession with his artificial creations? Or, is this currently a silent epidemic impacting computer scientists and others across the country, on a scale that’s

difficult to measure? There is evidence of the latter. For example, Google engineer Blake Lemoine was fired by the company in 2022 for going public with the belief that he had created a digital being, and he argued that he was not just anthropomorphizing a language simulator. Lemoine is also an ordained Christian priest, so feeling like God was perhaps easier for him to understand than my academic peer, even if it cost him his job.

Our society may be at the precipice of a whole new kind of mental health crisis. Call it A.I. psychosis, as people have already died from interacting too deeply with these algorithms, such as the Belgian man who committed suicide because an app called Chai told him that killing himself would help the environment. Yet, in a hilariously dark twist, many therapists are also turning to A.I. as a method of treating mental illness. We will live in a world where you can be driven to madness by your chatbot, and then your human doctor can prescribe a chatbot to help you.

We really are living in a Philip K. Dick novel!

The obvious threats these generative programs pose to our society have warranted calls for a slowdown by many experts in the field, such as Apple co-founder Steve Wozniak and public intellectual Yuval Noah Harari. Yet big tech companies like Google and Microsoft have fired their A.I. ethics teams for trying to implement policies to this effect. It doesn't matter how many people are negatively affected by these programs—they are simply the casualties of progress.

Grinding up humans for profits is nothing new: it is a central feature of capitalism. The difference this time is that it isn't just enslaved peoples or workers being thrown into the money-making furnace. As evidenced by my fellow graduate student's mental breakdown, even our best and brightest can be sacrificed to our new robot overlords. The A.I. tide is out at the beach and the tsunami is en route, so what do we do?

Looking back at the history of oppression, new problems need to first be named. Feminist theory taught us that women being treated as objects in the workplace was sexual harassment. Critical race theory informed us that white people having an inherent advantage that could never be obtained by non-white people was "whiteness as property." Users of these parrots are losing touch with reality, so A.I. psychosis is a tempting moniker. Like internet addiction, people are spiraling into their machine intelligences and anthropomorphizing them into human beings, and it will only get worse. Tech developer Enias Cailliau has even created GirlfriendGPT, a companion simulator that will send you voice notes and selfies.

No matter what information is presented to those afflicted with A.I. psychosis, they will insist that their parrots are real people. Like other mental health issues, our solutions to breaking people out of this cycle are similar. Aversion therapy with a negative stimulus whenever they interact with a chatbot. Exposure therapy to glitches/breaks in the A.I. behaviors to prove that these digital personalities aren't real. And, most importantly, nature therapy to get them away from the screens which continue to destroy their lives. Go touch some grass, as they say.

Reality is not the same as science fiction, even though we are living in a science fictional world. We don't need to worry about androids being able to pass the Voight-Kampff empathy test from *Bladerunner* anytime soon—it is people who are transforming themselves into machines.

Jess Flarity is a PhD candidate at the University of New Hampshire studying science fiction and gender.

fifth Estate

Jess Flarity
A.I. Psychosis & Personality Simulators
“How do I know you’re a human?”
2023

<https://www.fifthestate.org/archive/414-fall-2023/a-i-psychosis-personality-simulators>
Fifth Estate #414, Fall 2023

fifthestate.anarchistlibraries.net