



Everyone Dies?

February 2, 2026

After Friday, when I [worried about robots taking over](#), I was glad to read a debunking of the AI Will Destroy Us All meme, so in vogue. In “[Superintelligent AI Is Not Coming To Kill You](#),” from the March issue of *Reason*, Neil Chilson argues that we shouldn’t freak out.

Not only do I not want to freak out, I don’t want to use AI very much—though I understand that, these days, sometimes it makes sense to consult the Oracles.

Chilson is reviewing a new book, [If Anyone Builds It, Everyone Dies: Why Superhuman AI Would Kill Us All](#), by Eliezer Yudkowsky and Nate Soares, who argue that “artificial intelligence research will inevitably produce superintelligent machines and these machines will inevitably kill everyone.”

Just like I feared on Friday!

Where the authors go wrong, Chilson argues, is that by “defining intelligence as the ability to *predict* and *steer* the world, Yudkowsky and Soares collapse two distinct capacities—understanding and acting—into one concept. This builds their conclusion into their premise. If intelligence inherently includes steering, then any sufficiently intelligent system is, by definition, a world-shaping agent. The alignment problem becomes not a hypothesis about how certain AI architectures might behave but a tautology about how all intelligent systems must behave.”

Today’s AI’s are “fundamentally about prediction. They predict the next element in a sequence.”

They aren’t necessarily taking action.

I hope Chilson’s critique holds true.

But we’ve caught AI lying, “just making stuff up”—though considering the nature of “Large Language Models” (the method by which modern AI works), “lying” may be the wrong word. Still, it just seems to me that at some point somebody’s—*everybody’s!*—gonna link the predictor to some sort of truly active mechanism.

Like street-ready robots.

This is Common Sense. I’m Paul Jacob.