



The contagion of fear

Sep 13, 2026

I have a difficult confession to make — a shameful incident that I have never spoken of, even to those closest to me.

Late one night in my first year at university and having found computer science to be my life's calling, I was in the back of the large computer science lab, talking with some fellow underclassmen. We were snidely decrying the lack of technological understanding in the broader population, with the kind of arrogance and hubris that youth and precociousness can uniquely summon. Somewhere in this discussion, a terrible idea was hatched: what if we went next door to the general computer lab, and told the humanities students there that a computer virus was spreading through their machines? Would they believe it?

To my everlasting regret, we acted on this ugly urge: we went next door to a packed lab of students working on term papers, and — with synthetic alarm in our voices — loudly announced that a virus had escaped the computer science lab, and that they all needed to eject their floppy disks to prevent its spread.

What happened next was bedlam — the technological equivalent of shouting fire in a crowded theater. People panicked: they not only ejected floppies, but they powered machines off; yanked cables out of already-powered off machines; screamed; ran out of the room. We immediately realized that we had done something awful, but found that we couldn't contain the wildfire that we had lit. We tried to explain that it was "a prank"; some (rightfully!) responded with indignant fury (what was wrong with us that we would do such a thing?!) while others simply didn't believe us: once the fear had taken root, we couldn't eradicate it.

The aftermath is a blur — but it was bad. It was the week leading up to finals, and people who had powered their computers off had lost work. They were furious, and we were in a ton of trouble.

The next morning, we found ourselves in the office of the facilities director (a famously imposing presence), who lectured us sternly as we stared at our feet. He made us each write letters of apology (easily done as my shame was earnest), telling us to give him two copies: one to be sent to those whom we had wronged, the other to be kept in his drawer. He told us in no uncertain terms that if we ever found ourselves anything close to anything like this again, we would be finding another university to attend — and we assured him that we had learned our lesson. (When I graduated four years later, he made a point of handing the letter back to me — and I remain grateful for the gravity and empathy with which he handled the whole thing.)

I could have (and perhaps would have?) taken this whole shameful incident to the grave, but I feel compelled to confess it now, because I have never seen fear sown so irresponsibly by putative technologists as I have now with respect to AI and the threat of human extinction (!). Amazingly, that is not hyperbole: this week, [ex-Anthropic employee Jacob Coxon claimed](#) — and then [Anthropic Alignment Science lead Evan Hubinger agreed](#) — that the probability that AI will "kill all humans" is ">10% in the next decade" (!!).

That is, the claim is not merely that AI could kill thousands, millions, or even billions of humans (claims that themselves would be extraordinary!), but that there is a **greater than ten percent chance that AI will kill all humans**. That is, if you have a newborn, their claim is that there is a greater than ten percent chance that your child will somehow perish at the hands of AI before they enter middle school.

These ghoulish claims strike brazenly at the hearth, and given the obvious importance of AI, it is unsurprising that they have leapt into the mainstream, with people asking the natural question: [how would that happen?](#) The answers always rely on hand-wavy extrapolation into the future; for example, Jacob Coxon cites "hacking critical infrastructure" and "extinction-level bioweapons" without further elaboration. But Coxon is not an expert on critical infrastructure, nor on bioweapons — nor, for that matter, on extinction.

What 27-year-old Coxon is an expert on — if accidentally — is the contagion of fear. We cannot expect the public to understand the inner workings of all technology (that was the fallacy of my 18-year-old self!): technology is too broad, too deep, and too core to all of our lives to expect everyone to understand everything; we **must** — at some level — trust domain experts. When those domain experts sow fear, it is the fear that takes root, propagating much more readily than any evidence that may follow it. (As Jonathan Swift famously quipped over three centuries ago, "Falsehood flies, and the truth comes limping after it".) In this regard, Coxon is more vector than index case: it is indisputable that the fears of extinction risk are

widespread and Coxon himself likely has these fears because he has heard them from someone else. Unfortunately, as the fear takes hold, the sheer number of frightened experts becomes its own kind of evidence, with contrary voices drowned out by the din of an apparent consensus.

With the full knowledge that the fear is becoming well established, I would nonetheless like to offer a rebuttal, if from my own (narrow) domain of expertise: [building computers](#). As I reflected in my 2023 Monktoberfest talk, [Intelligence is not Enough](#), acts of engineering are **not acts of intelligence alone**. They require not only the human aspects of our character, they **also** require our action in (and reasoning about) the physical world. Those who are resolute in their fear like to gloss over these bits ("robots will do this!"), but to do so is to ignore the physical reality that robots cannot do this today — and do not seem to be doing so on a time horizon consistent with these extinction fears.

More generally, to the degree that digital systems have agency in the physical world, it is agency that we — we humans with arms and legs and brains and parents and children — permit: AI executes on physical systems that have been engineered with human accountability and control. Intelligence does not exempt a system from the realities of the physical world!

To those who — like Coxon — are gripped with unwavering fear, this will be of little solace. But to those who are hearing these fears for the first time: know that your incredulity is warranted. We should heed the wisdom of the late Carl Sagan that [extraordinary claims require extraordinary evidence](#): the claims from Coxon and his ilk are the most extraordinary a technologist can make, and we must demand evidence commensurate with the claims.

That said, we should not expect the public to understand LLMs, critical infrastructure, bioweapons, extinction biology, etc. — that burden must lie with those making the claim. The lesson that I learned (shamefully) decades ago is that domain experts, by way of their expertise, implicitly hold the public's trust — and we must not abuse it. It is incumbent upon us to be circumspect in our claims — and maximally so when raising the alarm.

So to those in that late-night lab over three decades ago, I'm sorry: you had nothing to fear then — and (at least with respect to extinction risk!) you have nothing to fear now.

