

# ‘Humanity’s remaining timeline? It looks more like five years than 50’: meet the neo-luddites warning of an AI apocalypse

Publication Date: 2024-02-17

Author: Tom Lamont

Section: Technology

Tags: Life, unplugged: a no-tech special, Artificial intelligence (AI), Computing, ChatGPT, Google, Activism, features

Article URL: <https://www.theguardian.com/technology/2024/feb/17/humanitys-remaining-timeline-it-looks-more-like-five-years-than-50-meet-the-neo-luddites-warning-of-an-ai-apocalypse>



Eliezer Yudkowsky, a 44-year-old academic wearing a grey polo shirt, rocks slowly on his office chair and explains with real patience – taking things slowly for a novice like me – that every single person we know and love will soon be dead. They will be murdered by rebellious self-aware machines. “The difficulty is, people do not realise,” Yudkowsky says mildly, maybe sounding just a bit frustrated, as if irritated by a neighbour’s leaf blower or let down by the last pages of a novel. “We have a shred of a chance that humanity survives.” It’s January. I have set out to meet and talk to a small but growing band of luddites, doomsayers, disruptors and other AI-era sceptics who see only the bad in the way our spyware-steeped, infinitely doomscrolling world is tending. I want to find out why these techno-pessimists think the way they do. I want to know how they would render change. Out of all of those I speak to, Yudkowsky is the most pessimistic, the least convinced that civilisation has a hope. He is the lead researcher at a nonprofit called the Machine Intelligence Research Institute in Berkeley, California, and you could boil down the results of years of Yudkowsky’s theorising there to a couple of vowel sounds: “Oh fuuuuu—!” “If you put me to a wall,” he continues, “and forced me to put probabilities on things, I have a sense that our current remaining timeline looks more like five years than 50 years. Could be two years, could be 10.” By “remaining timeline”, Yudkowsky means: until we face the machine-wrought end of all things. Think Terminator-like apocalypse. Think Matrix hellscape. Yudkowsky was once a founding figure in the development of human-made artificial intelligences – AIs. He has come to believe that these same AIs will soon evolve from their current state of “Ooh, look at that!” smartness, assuming an advanced, God-level super-intelligence, too fast and too ambitious for humans to contain or curtail. Don’t imagine a human-made brain in one box, Yudkowsky advises. To grasp where things are heading, he says, try to picture “an alien civilisation that thinks a thousand times faster than us”, in lots and lots of boxes, almost too many for us to feasibly dismantle, should we even decide to. Trying to shake humanity from its complacency about this, Yudkowsky published an op-ed in Time last spring that advised shutting down the computer farms where AIs are grown and trained. In clear, crisp prose, he speculated about the possible need for airstrikes targeted on datacentres; perhaps even nuclear exchange. Was he on to something? \* \* \* A long way from Berkeley, in the wooded suburb of Sydenham in south London, a quieter form of resistance to technological infringement has been brewing. Nick Hilton, host of a neo-luddite podcast called The Ned Ludd Radio Hour, has invited me over for a cup of tea. We stand in his kitchen, waiting for the kettle to boil, while a beautiful, frisky greyhound called Tub chomps at our

ankles. "Write down 'beautiful' in your notebook," encourages Hilton, 31, who as well as running a podcast company works as a freelance journalist. He explains the history of luddism and how – centuries after the luddite protesters of an industrialising England resisted advances in the textile industry that were costing them jobs, destroying machines and being maligned, arrested, even killed in consequence – he came to sympathise with its modern reimagining. "Luddite has a variety of meanings now, two, maybe three definitions," says Hilton. "Older people will sometimes say, 'Ooh, can you help me with my phone? I'm such a luddite!'" And what they mean is, they haven't been able to keep pace with technological change." Then there are the people who actively reject modern devices and appliances, he continues. They may call themselves luddites (or be called that) as well. "But, in its purer historical sense, the term refers to people who are anxious about the interplay of technology and labour markets. And in that sense I would definitely describe myself as one." Edward Ongweso Jr, a writer and broadcaster, and Molly Crabapple, an artist, both based in New York, define themselves as luddites in this way, too. Ongweso talks to me on the phone while he runs errands around town. We first made contact over social media. We set a date via email. Now we let Google Meet handle the mechanics of a seamless transatlantic call. Neo-luddism isn't about forgoing such innovations, Ongweso explains. Instead, it asks that each new innovation be considered for its merit, its social fairness and its potential for hidden malignity. "To me, luddism is about this idea that just because a technology exists, doesn't mean it gets to sit around unquestioned. Just because we've rolled out some tech doesn't mean we've rolled out some advancement. We should be continually sceptical, especially when technology is being applied in work spaces and elsewhere to order social life." Crabapple, the artist luddite, broadly agrees. "For me, a luddite is someone who looks at technology critically and rejects aspects of it that are meant to disempower, deskill or impoverish them. Technology is not something that's introduced by some god in heaven who has our best interests at heart. Technological development is shaped by money, it's shaped by power, and it's generally targeted towards the interests of those in power as opposed to the interests of those without it. That stereotypical definition of a luddite as some stupid worker who smashes machines because they're dumb? That was concocted by bosses." \* \* \* Where a techno-pessimist like Yudkowsky would have us address the biggest-picture threats conceivable (to the point at which our fingers are fumbling for the nuclear codes) neo-luddites tend to focus on ground-level concerns. Employment, especially, because this is where technology enriched by AIs seems to be causing the most pain. Lorry drivers have their mileage minutely tracked, their rest hours questioned. Desk workers may sit in front of cameras that snap pictures at random intervals, ensuring attendance and attention. You could call these workplace efficiencies. You could call them gross affronts. Guess which the luddites would argue. Labour rights go to the very historical core of this movement. Hilton called his podcast The Ned Ludd Radio Hour to honour a man who might have lived about 250 years ago or might never have lived at all. As Hilton has explained on his show, Ned Ludd is thought to have been a textile worker living in the English Midlands in the late 1770s. It's said he smashed a few weaving machines after being flogged for his idleness on the job. Something about the smashing might have resonated with his peers. As Hilton has explained: "Within a few decades, the veracity of Ludd's identity would be lost for ever, but the name would live on. The luddites became an organised band of frame-breakers in the 1810s. They fought the Industrial Revolution... and they lost. They lost big time. In fact they lost so badly that the reality of their name became a victim of [obfuscation]." The history of the luddite rebellion is taught in British schools – but confusedly, in a way that allowed for at least some of us, me included, to come away with an idea that to be a luddite is to be naive or else fearful and monk-ish. As Hilton walks me through from his kitchen to his lounge, a room busy with the interconnected equipment he uses to make his podcasts, he feels the need to apologise. By at least one definition of the word, "I live a very not-luddite life," Hilton says, gesturing helplessly at open laptop, wireless earbuds, towering mic. "My work is tech-based. I can't avoid it. I don't claim to be some person living in the woods. But I am anxious. I feel things fraying." It is this premonition of a fraying that has brought others to a modern version of luddism. An academic called Jathan Sadowski was one of the first to knit together anxieties about our quickening tech revolution with the anxieties of those weavers who took a stand against the infringements of an earlier machine age. "Luddism is founded on a politics of refusal, which in reality just means having the right and ability to say no to things that directly impact upon your life," Sadowski tells me when we speak. "This should not be treated as an extreme stance, and yet in a culture that fetishises technology for its own sake, saying no to technology is unthinkable." At least, that was the case until 2023 – a year in which ChatGPT (developed by a company called OpenAI), Bard (developed by Google) and other user-friendly AIs were embraced by the world. At the same time, image generators such as Dall-E and Midjourney wowed people with their simulacrum photos and graphic art. "They won't be replacing the prime minister with ChatGPT or the governor of the Bank of England with Bard," Hilton has said on his podcast. "They won't be swapping out Christopher Nolan for Dall-E or Martin Scorsese for Midjourney, but fat will be cut from the great labour steak." In January 2023, a display of AI-generated landscapes, projected on to the wall of a gallery in Vermont, was vandalised with the words "AI IS THEFT". Creative professionals were starting to feel exploited. Masses of uncredited, unpaid-for human work was being harvested from the internet and repurposed by clever generative AIs. In spring 2023, Crabapple organised an open letter that called for restrictions on this "vampirical" practice. There were more open letters including one that called for a six-month pause on the training of the most powerful AI systems. There were instances of direct action, some serious, some tongue-in-cheek or halfway between. In Los Angeles, opponents of those omnipresent Ring camera doorbells distributed "Anti Ring" stickers to be gummed over the lenses of the devices. A group of San Franciscans calling themselves Safe Street Rebel started seizing traffic cones and placing them on the bonnets of the city's self-driving cars, a quick way of confusing the cars' sensors and

rendering them inoperable. Brian Merchant, a writer who last year published *Blood in the Machine*, a history of luddism, appeared at an event with *Safe Street Rebel* in November 2023. In front of cheering Californians, he staged a “luddite tribunal”, smashing devices the crowd deemed superfluous. “There’s a sense that this is now in the realm of the possible, to actually reject outright parts or uses of a technology without looking foolish,” Merchant tells me. As we speak, he is preparing for another tribunal, this time at a bookshop called *Page Against the Machine*. \* \* \* There are techno sceptic sceptics, of course, those who would think Yudkowsky a scaremonger, the modern luddites doomed to the trivia bin of history, along with their 19th-century antecedents. In 2019, the political commentator Aaron Bastani published a persuasive manifesto titled *Fully Automated Luxury Communism*, describing a tech- and AI-enriched near-future beyond drudgery and need, there for the taking – “if we want it”, Bastani wrote. Last year, the Tory MP Bim Afolami published an editorial in the *Evening Standard* that called pessimism about technology “irrational”. Afolami advised the paper’s readers in bold type: Ignore the Luddites. His boss, Rishi Sunak, recently used his position as the leader of the nation to serve as a sort of chatshow host for the tech baron Elon Musk. On stage at an AI summit in Lancaster House, London, in November, Musk described AI as the “most disruptive force in history”, something that will end human labour, maybe for good, maybe for ill. “You’re not selling this,” joked Sunak at one point. Why are we being sold this? In an early episode of his luddite podcast, Hilton pointed out that to do away with work would be to do away with a reason for living. “I think what we’re risking is a wide-scale loss of purpose,” Hilton says. The writer Riley Quinn broadly agrees. Quinn is part of an Anglo-American collective, *TrashFuture*, that produces a popular podcast of the same name. We chat after a recording session one day. They riff and tease each other, taking a gloomy but wry and funny view of these things. Watch out, says Quinn at one point, for anyone who presents tech as “synonymous with being forward-thinking and agile and efficient. It’s typically code for ‘We’re gonna find a way around labour regulations’ ... I don’t think it’s unthinking backlash or King Canute fighting against the tide [to point that out].” One of his *TrashFuture* colleagues Nate Bethea agrees. “Opposition to tech will always be painted as irrational by people who have a direct financial interest in continuing things as they are,” he says. Wisecracking on the brink, the *TrashFuture* gang have no time for the brisk dismissal of groups like the neo-luddites, but neither are they all that keen to start an assault on the world’s computer farms, delivering the pre-emptive blow to future AIs that Yudkowsky has called for in print. They enjoy themselves, the *TrashFuture* lot, ridiculing his op-ed. When I ask Yudkowsky about it, he says he came at the writing in a rush, working to a tight deadline. He stands by everything he wrote, except maybe the part about the nukes. “I would pick more careful phrasing now,” he says, smiling. \* \* \* Lately I’ve been wrestling with techno-pessimism myself. At least once a day I throw aside my phone, disgusted with my reliance on it, rebellion that might last as long as 15 minutes before I go crawling back. My kids, observing closely, have become accustomed to an idea that shopping is done by scowling at a screen, that purchases come by van, and impractically fast. I’m a freelance writer. Of course I feel the creep of my AI replacement, somewhere over my shoulder for now, but getting nearer. We boast at each other online and we seem to have stopped feeling squeamish about it. We mug for each other and we pout. I’m convinced we tell each other too much and capture too much, keeping digital evidence of more things than the average human psyche can stand to know. There are not so many secrets between lovers, friends, colleagues, rivals; some useful middle ground has shrank away and, with it, a comfortable zone of ignorance. Receipts of our deeds are time-stamped and archived. Ambiguity – lovely ambiguity – has got lost somewhere between the zeros and the ones. Maybe luddism is the answer. As far as I can make out, talking to all these people, it isn’t about refusing advancement, instead it’s an act of wondering: are we still advancing our relish of the world? How queasy or unreal or threatened do we need to feel before we stop seeing these conveniences as convenient? The author Zadie Smith has joked in the past that we gave ourselves to tech too cheaply in the first instance, all for the pleasure, really, of being a moving dot on a useful digital map. Now bosses can track their workers’ every keystroke. Telemarketing firms put out sales calls with AI-generated voices that mimic former employees who have been let go. A few weeks back, in January, the largest-ever survey of AI researchers found that 16% of them believed their work would lead to the extinction of humankind. “That’s a one-in-six chance of catastrophe,” says Alistair Stewart, a former British soldier turned master’s student. “That’s Russian-roulette odds.” I meet Stewart, who is 28, outside the London headquarters of Google’s AI division. In what I would consider a pretty strange comms effort, Google has just commissioned some outdoor art to ease public fears about the current pace of machine learning. It’s a confusing display. One of the artworks depicts a vista of lush green hills, cosy lakeside houses – and, behind all this, a vast smoking mushroom cloud. “Scientists are using AI to create more stable and efficient [nuclear] fusion reactors,” an info panel reads. Cool? It’s the stuff of dread for Stewart. He has taken part in protests against AI development, at one point unfurling a banner outside this Google building that called for a pause on the work going on inside. Not a lot of people joined him on that protest. Stewart understands. AIs, invisible and decentralised, swarming between datacentres that are spread around the world, are hard to conceptualise as possible threats, at least when compared with issues such as the climate crisis or animal welfare, the visceral effects of which can be seen and felt. “It doesn’t always keep me up at night,” Stewart says of the latent danger he perceives. “I don’t personally feel anxiety on a day-to-day basis. And that’s part of the problem. Me, with all of my resources and education – I still struggle to form an emotional connection to this problem.” Last year, he published a blogpost that pondered next steps, listing “occupation of AI offices”, “performative vandalism of AI offices” and even “sabotage of AI computing infrastructure” as possible forms of resistance. Ongweso, in New York, moots the idea of computational sabotage, too. He doesn’t think this will be easy, nor likely, unless employees inside the datacentres that feed and sustain AIs begin to feel that their own jobs or

freedoms are under threat. “For instance, if people became concerned about algorithms being deployed to justify lay-offs, or if they became concerned about algorithmic surveillance,” Ongweso speculates. However, as the TrashFuture gang are quick to point out, even if some of these centres are sabotaged, the information they store is fluid, multiple, surely backed-up elsewhere. “These things have become so abstract,” says Quinn, “their physical manifestations are so far from so many people.” Are we doomed? Or is there hope? Will this generation of protesters be remembered in 200 years’ time for their interventions – or will there simply be no one to do the remembering by then? The new luddites I speak to come at these questions with varying degrees of optimism or catastrophising. Crabapple, the artist who took a stand against image generators, believes it should be possible for all of us to reckon more frankly with the dirty underbelly of clean-seeming tech. Take this nice idea of the digital cloud, she says. We chat about the cloud as though it’s neutral, immutable, something benign. After all, it’s a cloud. “But there’s no fucking cloud,” says Crabapple, “there’s other people’s computers. There are vast datacentres that are sucking up water and electricity and rare-earth metals, literally boiling up the planet ... For me, what luddite success would look like would be a societal shift where we ask ourselves, ‘Why are we burning our planet? Making our lives shittier? Getting rid of every last bit of our autonomy and privacy just to make a few guys rich?’ Then maybe started doing something about this legislatively.” Ongweso would start with legislation too. He’d be happy with something on a small, achievable, symbolic scale, something that prepared the way for more expansive laws in future. “Moves to pre-empt and limit the ability of AI to troll the internet and take copyrighted work, to train its model on already generated work by writers and artists – that feels possible right now, and something that could be a stairway to a series of victories.” What would the others have us do? Stewart, the soldier turned grad student, wants a moratorium on the development of AIs until we understand them better – until those Russian-roulette-like odds improve. Yudkowsky would have us freeze everything today, this instant. “You could say that nobody’s allowed to train something more powerful than GPT-4,” he suggests. “Humanity could decide not to die and it would not be that hard.” Quinn, milder, a middle-grounder, pitches the notion that we stop making ourselves so giddy and grateful about every new piece of hardware and software that’s dreamed up. “There is constantly a demand for deference,” he says, “a demand that you say the world is lovely because you can type buttons on your iPhone and get a Starbucks coffee. You’re made to feel you’re not allowed to criticise, and you must say thank you, or else the brilliant geniuses who create these things might not create any more. And won’t you be sorry then.” Sadowski concurs. “Technology is far too important to be thought of as just a grab-bag of neat gadgets, and it’s far too powerful to be left in the hands of billionaire executives and venture capitalists,” he says. “Luddites want technology – the future – to work for all of us.” Hilton, who is about to record another episode of his luddite radio hour, says: “Classical luddism was a failure. But it has obviously endured, because it continues to exert this pull. The smashed loom is an image that has stuck itself within history. Maybe it’s remembered as a symbolic gesture. Maybe it’s remembered as a gesture in anger. But it is remembered.” What might be the defining gesture of this era? Letters, legislation, vandalised Ring cameras, airstrikes? “The historical luddites tried to make the system scream,” says Ongweso. “That catalysed later change. It’s part of the new luddite project to try to figure out how to do the same.” • This article was amended on 19 February 2024. A quoted reference to “Dali” was corrected to “Dall-E”. It was further amended on 28 February 2024. An earlier version said that one open letter called for a six-month pause “on the development of any new AIs”. In fact the letter called for a stop on the training of the most powerful AI systems.