

Opinion | This Is the Good Scenario?

Ezra Klein

Advertisement

[SKIP ADVERTISEMENT](#)

Ezra Klein

Oct. 18, 2026



Listen · 15:23 min

Sign up for Ezra Klein's weekly newsletter. A guide to what Ezra is reading and listening to.

Imagine Americans were offered this deal: Hit a big red button and China becomes the world's first hyperpower. It rockets past the United States in technical sophistication, manufacturing capacity and wealth generation. China becomes so technologically advanced, and its human and robotic work force so huge, that Americans are left with only the dimmest comprehension of what is playing out on the economic, scientific and perhaps even geopolitical frontier.

But the rewards for disempowering ourselves are magnificent. The global economy booms, and Americans grow richer than ever before. We are given new cures and cheap energy and wondrous entertainments. Machines perform our labor and mood-enhancing drugs protect our psyches. We need not struggle — perhaps we need not even work — to live lives, long lives, of luxury and ease.

There, then, is the trade: Cede control and much that we have struggled for can be ours. But control, once gone, is gone.

Would you press the button?

The A.I. conversation centers questions of catastrophe. I want to center the question of success: What if A.I. works not as its architects fear, but as its inventors and investors hope? Wander outside of Silicon Valley's p(doom) — their shorthand for the probability that A.I. extinguishes the human race — and you might notice that their p(utopia) has problems, too. They imagine something like the world I sketched above, but more extreme, with A.I. playing the role assigned to China.

An amusing collision of these visions came when Jacob Coxon, the 27-year-old OpenAI and Anthropic whistle-blower, [sat down](#) with Jon Stewart on “The Daily Show.” Coxon was there to warn that A.I. might kill us all, but when Stewart offhandedly mocked the “future utopia” of the A.I. architects, Coxon stopped him. “Well, I am a believer in the future utopia,” Coxon said. But what he said next sounded grimmer: “So I heard you talking about jobs. I think there will be no jobs.”

Coxon went on to describe “the world that is being envisioned by the people building this technology” as “no jobs. Universal high income. The dividends from entirely automated economic production can be distributed among real humans.”

This vision, or some variant of it, is common enough at A.I. companies. I have heard it often. But most people, face to face with this idea, are incredulous at hearing this world — in which virtually all advanced cognitive work is done by A.I. — described as utopian. Stewart certainly was. “So we’ll all just be performing in a play?” he asked.

I tend toward [skepticism](#) of the jobless future. Human beings are good at nothing if not finding something to do. (Who would have imagined, in our grueling agrarian past, that podcasting could be a full-time job?) But I find it easy to imagine how human control and comprehension of much of society could be handed over to

machines. A few recent developments — some reflecting the outer boundary of what advanced A.I. can do and the other reflecting its advancing utility in daily life — reveal the ways it is already happening.

A few weeks ago, OpenAI [dumped](#) more than 700 manuscripts proposing solutions to nearly 400 advanced mathematical problems all at once. If you're not a mathematician, and you are not close to any mathematicians, this may have passed by unnoticed. Within mathematics, it was an earthquake.

I am not a mathematician, but my father is. The area of math he worked in had to do with the way electrons move through disordered crystals. Sitting at the center of his field were two problems that were more than just unsolved; they were, my father says, “totally out of reach — nobody had any idea how to solve them.” OpenAI's systems solved them in an afternoon. “It is the end of mathematics as we know it,” he texted me. “It is the first time I am glad I have retired.”

Sign up for the Opinion Today newsletter Get expert analysis of the news and a guide to the big ideas shaping the world every weekday morning.

The newly formed Association for Human Mathematics, which has about 1,000 members, quickly released [a statement](#): “Mathematicians did not ask for this work to be done.” It's an interesting sentence. Should mathematicians be able to decide which math problems do and do not get solved? How do we weigh the benefits and costs of a world where the frontier of mathematics is advancing rapidly, but human beings no longer understand it?

Imagine the OpenAI process running repeatedly, over a period of years: A.I. swarms solving problems no human being has solved, then using those solutions to pose new problems, which are solved by other A.I. swarms that have absorbed the work done by their predecessors, which allows for the formulation of yet more problems, and on and on. It is too much new math for humans to comprehend, but A.I.s can absorb it easily, and they do, driving forward advances in material sciences and energy production. Human beings do not understand all this new science, but they do benefit from it.

Is that world so different from the one we now inhabit? Few of us understand much about the borderlands of human knowledge. I cannot remotely comprehend the math my father did. How different is it if none of us can understand the math the A.I. does? And yet, a world where the farthest edge of knowledge and progress rests in algorithmic systems, rather than human minds, does seem different to me. Human control — which is, in my view, inseparable from human comprehension — is worth something, is it not?

As I was writing this, I [read about](#) a study in which theoretical physicists posed 50 complex, unpublished problems to advanced A.I. systems. These problems were difficult but solvable, taking a physicist who was expert in the particular area of the problem about a day to solve. The best A.I. solved 98 percent of them, at a cost of less than \$3 per problem.

The physicists initially thought the A.I. had performed worse than it had, but the model insisted that it was the humans, not the A.I., that had erred, and it was right. “Frankly, I wasn't expecting this day to come as soon as it has,” Suvrat Raju, one of the physicists behind the study, [wrote](#). “After we found these results, I couldn't sleep properly for a couple of nights.”

The grimmest [A.I. stories](#) often center on a single moment when superintelligence slips out of our control and begins plotting our demise. But the scenario I find more likely and more irresistible is called “[gradual disempowerment](#).”

In this telling, there is no one moment when humanity loses control; there is instead a progression of small moments when we willingly give control away because it is the rational thing to do. Firms turn first more of their work and then more of their decision making over to advanced A.I.s because it gives them a competitive edge. Governments turn more battlefield decisions over to A.I. because it makes their forces swifter and more lethal. A.I. companies turn to recursive self-improvement — where the A.I. autonomously creates the next A.I. model — because it vaults them ahead of their competitors.

The key to this process is competitive pressure. Companies and governments and A.I. labs that resist are outcompeted by peers that give A.I.s more control and autonomy. Advances in A.I. systems — and, eventually, in A.I.-controlled drones and robotics — keep expanding the amount of work A.I.s can do better than human beings. Even assuming the A.I.s remain helpful and pliant, more and more of society’s basic functioning slips beyond the grasp of human beings.

Which brings me to a second, more mundane, story in A.I. Meta’s [Muse assistant](#) has topped the Apple app charts for weeks now. At its core, Muse is a fairly simple A.I. model. What makes it useful is that it is designed to operate with access to much more of your digital life — your email, your calendar, your social media, your bank accounts, your cloud storage — and then automatically take actions on your behalf.

I feel queasy about offering an A.I. so much access to my life, but my friends who use it tell me it has made them feel superhuman: They are more responsive and more responsible than they have ever been, because there is an always-on A.I. agent monitoring their digital life and catching problems and following up on emails and making reservations and reminding them of their priorities. The catch, of course, is that there is someone on the other side of all these A.I. actions.

Already, we’re [hearing](#) of businesses getting swamped by automated calls and emails. As Muse-like agents spread, those businesses will employ their own A.I.s to answer the influx. As A.I. agent use like this becomes more widespread, anyone who doesn’t use an agent may begin to fall behind. We will be less responsive than those who do use agents, or simply overwhelmed responding to other people’s agents.

Eventually it will be our A.I.s bargaining with other A.I.s, our A.I.s calling the restaurant’s A.I. for reservations, our A.I.s emailing our colleague’s A.I.s. The edge will go to those who grant them more intimate access or autonomous control. And the logic of competition will roll forward. What would seem bizarre one year will become commonplace the next. Much of our world will be turned over to A.I.s.

We have seen this before. There was a time when we could resist using email or smartphones, but for most of us, that time has passed. It is too much of a disadvantage in modern society for most of us to forgo these technologies, even if we lament their consequences in our lives. The same will likely prove true with A.I.

Perhaps this is just how it has always been. Isn’t the story of modern society human beings trading comprehension and control of their world for technological and economic advancement? I would not trade

places with my ancestors, even though society is less legible to me than it was to them. That air-conditioning or M.R.I. machines work is enough; I do not need to know, at a fundamental level, how they work. But that story has, until now, remained a fundamentally human story. That day may be ending.

It's easy enough to answer the question of whether we want A.I. that kills us all. We don't. But do we want A.I. that enriches us all, if the cost is our control and comprehension over the world in which we live? What is the value to human beings maintaining primacy over our own struggle, and how do we weigh it against the easing of that struggle? I want the cures and the theorems and even, on some level, the capable digital assistants. And I also believe the struggle for human progress is essential to the human story, and even the human soul.

I don't find these questions easy to answer. I don't even find them easy to clearly pose. But right now, we are on a path to letting them be answered for us, in countless small and mostly unnoticed moments when it will make competitive or personal or political sense for the humans to step back and let the machines step forward.

In his [much-misunderstood book](#) "The End of History and the Last Man," Francis Fukuyama posed a similar question. It seemed unlikely, Fukuyama thought, that any political or economic system would surpass capitalism and liberal democracy at driving forward material prosperity. The question was whether there were human drives that liberal democracy would not only fail to meet, but fundamentally betray.

Fukuyama thought there was: Man did not just want a paycheck and an apartment; he wanted recognition and struggle. He wanted to be seen as not just equal to other men in some spiritual or political sense, but as better than other men. The fight for recognition is a fundamental human drive. And the inability of technocratic, liberal societies to even comprehend that urge might destabilize them.

"We inhabitants of liberal democratic countries are by now so used to accounts of current events that reduce motivation to economic causes, so thoroughly bourgeois in our own perceptions, that we are frequently surprised to discover how totally noneconomic most political life is," Fukuyama wrote. "Indeed, we do not even have a common vocabulary for talking about the prideful and assertive side of human nature that is responsible for driving most wars and political conflicts."

A.I. seems, to me, to pose Fukuyama's question in an even sharper form. If it plays out as its proponents hope, it promises levels of prosperity and possibility conventional political systems never imagined. But it does so at the cost of betraying the human drive for recognition and struggle. Fukuyama worried about liberal democracy because its promise of equality, even when equality could be achieved, would not be enough. He saw liberal democracy as built on an incomplete conception of what human beings actually want. But A.I. goes further: It promises inferiority as a condition of prosperity.

Stewart saw this problem immediately. "If I may poke a hole in that utopia," he said to Coxon, "I think, for the most part, even when we're doing jobs we don't particularly like, that humans like to feel relevant When people my age retire, they often die, because they're no longer connected to purpose."

Coxon, to his credit, didn't dodge the question. Many people hate their jobs, he said. And I'd add: Many

people love retirement, though they loathe unemployment. But I don't think the problem here is best understood as jobs, or at least not solely as jobs. There is much human beings will prefer to do for and with each other even if A.I.s can do it better. That a machine can beat a chess grandmaster has not meant the end of human beings playing chess. But humanity's sense of purpose and control and primacy is a harder and more ineffable thing to replace.

"Once you began admitting explanations in terms of purpose — well, you didn't know what the result might be," Aldous Huxley writes in "Brave New World." The book imagines a world in which all desires are fulfilled but struggle is absent. In the society Huxley imagined, debates over purpose were considered subversive. They could leave people believing "that the purpose of life was not the maintenance of well-being, but some intensification and refining of consciousness, some enlargement of knowledge. Which was, the Controller reflected, quite possibly true. But not, in the present circumstance, admissible."

Liberal democracies are threatened today not because they failed to deliver prosperity that was unimaginable to our predecessors, but because they allowed so much inequality and alienation and powerlessness to fester amid that prosperity. We want wealth, but we want wealth in part because it promises to grant us dignity. You can see, in the fury of the politically radicalized rich, how insufficient wealth is when severed from the recognition they believed it would accord them.

I am not saying that A.I. will, or even can, create this world. But it is the world many of its architects envision. Their vision for A.I. is not merely that it replaces white-collar work but that it races ahead of our most brilliant minds, and with each day that passes, with each new model that is released, with each new piece of evidence that A.I. can solve problems we cannot, this becomes a little more plausible.

The people building A.I., Stewart said, "don't particularly seem like they have an understanding of the human condition." If given the chance, almost all of them would push the button. Would you?

Ezra Klein joined Opinion in 2021. He is the host of the podcast "The Ezra Klein Show" and the author of "Why We're Polarized" and, with Derek Thompson, "Abundance." Previously, he was the founder, editor in chief and then editor at large of Vox. Before that, he was a columnist and editor at The Washington Post, where he founded and led the Wonkblog vertical. He is [on Threads](#).

Advertisement

[SKIP ADVERTISEMENT](#)