

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

I'm Ezra Klein, this is the Ezra Klein Show.

Welcome to the Ask Me Anything episode.

I am your guest, Ezra Klein, here with Roje Karma, our senior editor, who is going to be asking me questions and proving that we are not, in fact, the same person, has sometimes been suspected.

But Roje, thank you for being here.

It's great to be here.

I appreciate you giving us the opportunity to prove once and for all we are indeed two different people.

So what you got for me?

So this AMA was a little bit unique in that we usually get a very wide range of questions without any particular subject dominating.

But this time we got absolutely flooded with questions on AI, you know, questions about existential risk and labor markets and utopias, and we're going to get to all of that.

But given how fast all of this is moving, I just wanted to start by checking in on where your head is right now.

So a couple of questions first, how are you thinking about AI at the moment?

And then second, what has your approach been to covering AI both in your writing and on the show?

I think as we go through questions, people are going to get a sense of how I'm thinking about it.

But I guess I'll say in the approach bucket, I am trying to remain open to how much I do not know and cannot predict.

Look, I enjoy covering things and have typically covered things where I think there is usually a body of empirical evidence where you can absorb enough of it to have a relatively solid view on where things are going.

And I don't think that is at all true with AI right now.

So here my thinking is evolving and changing faster than it normally does.

I am entertaining the simultaneous possibility of a more radical set of perspectives, everything from the existential risk perspectives of, and we can talk about this, that we will create non-aligned hyperintelligent computerized agents that could pose a catastrophic risk or create a catastrophe for humankind, all the way to in the 100, 200-year range, we could be in a post-recutopia, and then of course the much more modest and I think likelier and more normie views of we'll get a lot of disinformation.

We will have a faster pace of scientific discovery.

So there is such a, I think, a wide range of possibilities here that you have to just sit with a lot of things potentially being true all at once.

And you or I, I should say, have to be willing to be learning and wrong in public and not collapsing into the tendency that is attractive and I think sometimes correct on other issues to say, this is what I think is likeliest.

This is my best read of what's going on.

And as such, this is the lane I'm going to cover or the interpretation I'm going to advocate for.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

To use a computer science metaphor that some people may know, both mentally and then in the podcast and in the column professionally, I'm in much more of an explore mode than an exploit mode.

So there's a lot to dig into from that answer and luckily we have a lot of great audience questions to get into some elements of it.

So you mentioned sort of the existential risk scenarios and we had a few different questions around those.

For instance, we have a question from Patrick A. who has been reading the arguments of folks like L.A. Zierdowski who are sounding the alarm on how A.I. is an existential threat to humanity.

And Patrick writes, quote, increasingly my feeling is that to the extent that the conversation around A.I. is fixated on the relatively short term, things like job loss, disinformation, bias algorithms, that as important as these issues are, we're whistling past the graveyard on this problem.

So what do you make of the most dire assessments of the risks posed by A.I. and what level of alarm do you feel about its dangers and why?

I have very complicated thoughts here.

Let me pick apart an idea in that question first and then I'll get to the bigger part of it.

So first, for people not fully familiar with this idea, existential risk around A.I. is fundamentally the prediction that there is at least a high probability that we will develop super intelligent artificial intelligences that will either destroy humanity entirely causing an extinction level event or displace us profoundly, disempower us profoundly in the way we have done to so many other animal species, so many of our forebears.

So that's the existential risk theory and we can talk about why people think that could happen and I will.

I am much less persuaded than some people that you want to disconnect that from medium term questions because when you're talking about existential risk, you're talking about what gets called a misaligned A.I.

So in one version of this, you have something that is all powerful but somewhat stupid. You build this hyperintelligence system and you say, you know, again, canonically here, make me paper clips and then it destroys the entire world and evades all of your efforts to shut it off because its one goal in life is to make paper clips.

It can make, you know, the iron in our bodies into a couple more paper clips on the margin and for that matter, it can, it needs to kill us because we might stop it from making more paper clips.

So that's one version and I think that sounds to people and they hear a little stupid.

It sounds to me when I hear it a little stupid because if you're going to make something that's smart, don't you think you're going to be able to program in some amount of reflectivity, some amount of uncertainty, some amount of, hey, check in with us, you know, about whether or not you're actually achieving the goals that we want you to achieve.

But there are of course more conceptually sophisticated versions of that.

So you are a rogue state, a North Korea, or maybe you're just an up and coming state and

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

you say, hey, I want you to make us as powerful as you can and you do not have the capabilities to correctly sort of program your system and the system causes a huge catastrophic event trying to destroy the rest of the power centers in the world.

You can really think about how that might happen.

I think the best way to approach this question though for me for right now is thinking about the ways you might get highly misaligned AIs and there are two versions of this.

One is what you might call the intelligence takeoff version, which is you get AI systems that become sort of smarter than human beings in a general intelligence way and way more capable of coding.

So now the AIs begin working on themselves is something I talked about with Kelsey Piper. The AIs create a kind of intelligence takeoff because in sort of recursively coding themselves to become better, they can go through many, many, many generations very, very quickly. And so you have this hyper accelerated evolutionary event which leads to in a fairly short time span, much smarter systems than we know how to deal with that we don't really understand their alignment and like all of a sudden we're out in the cold.

That could happen.

I have a lot of trouble knowing how to rate the possibility of rapid intelligence takeoff.

I have some points of skepticism around it, around how much for instance, how phenomenal of a capability set you can get from just say absorbing more and more online text, whether or not it's actually going to be true of these systems that have a lot of trouble distinguishing true things from false things are going to be able to so effectively improve themselves without creating a whole lot of problems.

But a version of this I find more convincing came from Dan Hendricks, who's a very eminent AI safety researcher.

He wrote a recent paper that you can find it's called natural selection favors AIs over humans.

And the point of his paper is I think it offers a more intuitive idea of how we can get into some real trouble whether or not you're thinking about existential risk or just a lot of risk. And he writes, as AI has become increasingly capable of operating without direct human oversight, as could one day be pulling high level strategic levers.

And if this happens, the direction of our future will be highly dependent on the nature of these AI agents.

And so you might say, well, look, why would we let them operate without direct human oversight? Why would we program these things and then turn over key parts of our society to them such that they could pose this kind of danger?

And what I appreciate about his paper is I think he gives a very, very realistic version of what that would look like.

So he writes that these AIs are basically going to get better, right?

That's already happening.

We can get that.

We're going to turn over things like make an advertising campaign or analyze this data set or I'm trying to make this strategic decision about what my country or my company should do.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

Look at all the data and advise me.

And he writes, eventually, AIs will be used to make the high level strategic decisions now reserved for CEOs or politicians.

At first, AIs will continue to do tasks they already assist people with, like writing emails. But as AIs improve, as people get used to them and it's staying competitive in the market demands using them, AIs will begin to make important decisions with very little oversight. And that to me is a key point.

What he is getting at here is that as these programs become better, there's going to be a market pressure for companies, potentially even countries to hand over more of their operations to them because you'll be able to move faster.

You'll be able to make more money with your high speed AI driven algorithmic trading.

You'll be able to outcompete other players in your industry.

Maybe you'll be able to outcompete other countries.

And so there'll be a competitive pressure where for a period of time, the institutions, companies, countries that do this will begin to prosper.

They will make more money, they will get more power, they will get more market power.

But that having done that, they will then have systems they understand less and less and have less and less oversight over holding more and more power.

And then he goes through a thing about why he thinks evolutionarily that would lead to selfish systems.

But the thing I want to point out about it is that rather than relying on a moment of intelligence takeoff, it relies on something we understand much better, which is it we have an alignment problem, not just between human beings and computer systems, but between human society and corporations, human society and governments, human society and institutions. And so the place where I am worried right now, the place where I think it is worth putting a lot of initial effort that I don't see as much of, is the question of how do you solve the alignment problem?

Not between an AI system we can't predict yet and humanity, though we should be working on that, but in the very near term, between the companies and countries that will run AI systems and humanity.

And I think we already see this happening.

Right now, AI development is being driven principally by the question of, can Microsoft beat Google to market?

What does meta think about all that?

So there's a competitive pressure between countries.

Then there is a lot of, you know, US versus China and other countries are eventually going to get into that in a bigger way.

And so where I come down right now on kind of existential risk is that when I think about the likely ways we develop these systems that we then create such that we have very little control over them, I think the likeliest failure mode right now is coming from human beings.

So you need coordinating institutions, regulations, governance bodies, etc. that are actually thinking about this from a broader perspective.

And I worry sometimes that the way the existential risk conversation goes, it frames this almost

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

entirely as a technical problem when it isn't.

It's at least for a while a coordination problem.

And if we get the coordination problems right, we're going to have a lot more leverage on the technical problems.

I think a lot of other people share a similar concern about these sort of incentives, about the speed at which everything is moving.

And one of the responses in the last week has been this open letter from more than a thousand, you know, tech and AI leaders, including simply high profile people like Elon Musk, along with like a lot of AI researchers.

And this letter was calling for a six month pause on any AI development more advanced than GPT-4.

And I think the concern that that letter comes from is the same concern you just outlined, right?

This is all moving too fast.

We don't like the incentives at play.

We don't know what we're creating or how to regulate it.

So we need to slow this all down to give us time to think, to reflect.

And so I'm wondering what you think of that effort and whether you think it's the right approach.

I think there's a good amount to be said for it.

But my thinking on this has evolved a bit.

So in my column here a few weeks ago now feels like months ago, maybe it was a month ago, but time is moving quite quickly.

But that this changes everything column where I released it on the podcast under my own view on AI or my view on AI.

I do end up calling that either for an acceleration of human adaptation to this or a slowdown in development.

I have become increasingly skeptical, whatever you think of the merits of a slowdown though, that it is a winning political position that is even a viable political position to make the two sides of this.

People who think AI is kind of cool and they enjoy asking their chatbot questions and they want the better help on term papers and marketing support and more immersive and relational porn that's going to come out and the sort of human to AI companions, all the things that are near term values here and all the things that companies want to make money on, that's all on one side, right?

Everybody who either doesn't care about AI or wants something from it.

And on the other side, just AI is bad.

Stop it, I'm scared.

And I don't mean that to dismiss that position because I think AI might be bad and at times I am scared.

But I think you actually need a positive view, much more so than people have of what you're doing.

And if you do a pause and you don't know what you're doing with that pause, right?

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

If that pause takes place, you do a six month stop.

And what the ideas of the people in the AI companies and in academia are going to try to spend six months more on interpretability?

What are the systems under which we have public input here?

How are we coming up with an agenda, right?

What are you doing with that time?

Otherwise you just delayed whatever is going to happen six months and maybe gave worse actors out there.

Although I want to be careful of this because I think people here at China, and I am not sure China is actually a worse actor on this than we are right now, because I think we are actually moving a lot faster and we are using the specter of China to absolve ourselves of having to think about that at all.

But putting that aside, because I also don't want China to have AI dominance for a bunch of very obvious reasons, I think.

I do think I am more inclined to say that what I want is first a public vision for AI.

I think this is frankly too important to leave to the market.

I do not want the shape and pace of AI development to be decided by the competitive pressures between functionally three firms.

I think the idea that we're going to leave is to Google Meta and Microsoft is a kind of lunacy, like a societal lunacy.

The one, I think that more than I understand this is slowing it down and I understand it is shaping technology.

And there are things that I want to see in AI.

I want to see a higher level of interpretability.

And when I say interpretability, I mean the ability to understand what the AI system is doing when it is making decisions or when it is drawing correlations or when it is answering a question.

When you ask JetGPT to summarize the evidence on whether or not raising wages reduces healthcare

spending, it will give you something, but we don't really know why or how.

So basically, if you try to spit out what the system is doing, you get a complete and comprehensible series of calculations.

There is work happening on trying to make this more interpretable, trying to figure out from where in the system a particular answer is coming from for trying to make it show more of its work.

But that work, that effort is way, way, way behind where the learning systems are right now.

So we've gotten way better at getting an output from the system than we are at understanding what the inputs were that went into it, or at least what the sort of mid-level calculations were that went into it.

One thing I would do, and I don't know exactly how to phrase this because I'm not myself an AI researcher, but I think that particularly as these systems become more powerful, we should have some level of understanding of what's going on in them.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

And if you cannot do that, then you cannot release it.

And so I think one totally valid thing to say, because this would slow down AI research or at least AI development, but it would do so for a cause, is to say that if you want to create anything more powerful than GPT-4 that has a larger training run or training set and is using more GPU power and all the rest of it, more compute power, then we want these levels hit for interpretability.

And it is not our problem to figure out how to hit it.

It is your problem.

But yeah, there's a lot of money here.

Start putting that money towards solving this problem.

There's a lot of places in the economy where what the regulators say is that you cannot release this unless you can prove to us it is safe.

Not that I have to prove to you that you can make it safe for me.

Like if you want to release GPT-5 and GPT-7 and Claude++ and whatever, you need to put these things in where we can verify that it is a safe thing to release.

And doing that would slow all this down.

It would be hard.

There are parts of it that may not be possible.

I don't know what level of interpretability is truly even possible here.

But I think that is the kind of thing where I want to say, I'm not trying to slow this down.

I'm trying to improve it.

I'm trying to make it better.

And by the way, that might be true even from the perspective of somebody who wants to see the AIs everywhere, it's only going to take one of these systems causing some kind of catastrophe that people didn't expect for a regulatory hammer to come down so hard it might break the entire industry.

If you get a couple of people killed by an AI system, for reasons we can't even explain, you think that's going to be good for releasing future AI systems because I don't.

There's one reason we don't have driverless cars like actually all over the road yet.

That's one thing.

Another is that these systems are being shaped in the direction they're being constructed to solve problems that are in the direction of profit.

So there are many different kinds of AI systems you can create directed at many different purposes.

The reason that what we're ending up seeing is a lot of chatbots, a lot of systems designed to fool human beings into feeling like they're talking to something human is because that appears to people to be where the money is.

You can imagine the money in AI companions and there are startups like Replica trying to do that.

You can imagine the money in mimicking human beings when you're writing up a Word document or a college application essay or creating marketing decks or whatever it might be.

But so I don't know that I think that's great actually.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

There are a lot of purposes you could turn these systems to that might be more exciting for the public.

I mean, it is still, to me, the most impressive thing AI has done is solve the protein folding problem.

That was a program created by DeepMind.

What if you had a prizes system where we had 15 or 20 scientific and medical innovations we want, problems we want to see solved by whatever means you can do it?

And we think these are big enough that if you do it, you get a billion dollars, right?

We've thought about prizes in other areas, but let's put them into things the society really cares about.

Maybe that would lead more AI systems to be tuned in the direction, not of fooling human beings into thinking they're human, but into solving important mathematical problems or into speeding up whatever it is we might want to speed up this kind of drug development.

So that's another place where I think the goals we actually have publicly, how you can make money off of this, I would like to see some real regulation here.

I don't think you should be able to make money just flatly by using an AI system to manipulate behavior to get people to buy things.

I think that should be illegal.

I don't think you should be able to feed into it surveillance capitalism data.

Get it to know people better and then try to influence their behavior for profit.

I don't think you should be allowed to do that.

Now you might want to think about what that regulation actually reads like in practice because I can think of holes in that.

But whether I'm right or wrong about those things, these questions should be answered.

And at the end of that answering process, I think is not a vision of less AI or more AI, but a vision of what we want from this technology as a public.

And one thing that worries me is that just the negative vision, let's pause it.

It's terrifying.

It's scary.

I don't think that's going to be that strong.

And another thing that worries me is that Washington is going to fight the last war and try to treat this like it was social media.

We wish we had had somewhat better privacy protections.

We wish we had had somewhat better liability maybe around disinformation, something like that.

But I don't think just regulating the harms around the edges here.

And I don't think just slowing it down a little bit is enough.

I think you have to actually ask as a society, what are you trying to achieve?

What do you want from this technology?

If the only question here is what does Microsoft want from the technology or Google, that's stupid.

That is us abdicating what we actually need to do.

So I'm not against a pause, but I am also not for pause being the message.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

I think that there has not been nearly enough work done on a positive public vision of AI, how it is run, what it includes, how the technology is shaped, and to what ends we're willing to see it turned.

That's what I want to see.

So let's dig in more into what that positive vision could be, because I think a lot of people hear about all of these risks, some of these existential scenarios, and their responses like, well, why should we be doing this at all, but we actually got some questions from audience members about sort of the possibilities that AI technology can unlock. And so, for example, Catherine E asks, while AI researchers think that there's a 10% chance of terrible outcomes, they think there's an even higher chance of amazing utopian outcomes. And here she's referencing a recent survey of leading AI experts that we can link to in the show notes.

So she continues, you've discussed the possible nightmare scenarios for AI, but do you see potential upsides?

Do you think your kids' lives might be better because of AI, not worse?

Yeah, I do think there's a lot of possibility here for good outcomes.

And I do think probably the good outcomes, or at least like the weird and mixed outcomes are a lot likelier than the totally catastrophic ones.

I will be honest that I ask this question of a lot of people in the space, and I don't find the answers I get are that good.

So I think the most common answer I hear is AI could become an unbelievable scientific accelerant.

And maybe, maybe, absolutely maybe.

I think the reason I've always been a little more skeptical of that than some people in the space is that while it's clearly true there are many scientific advances you could make just by being a hyperintelligent thing reading papers.

I mean, Einstein was not running direct experiments, he was creating brilliant thought experiments that led to, over time, tremendous revolutions also in industry and technology and so on.

I do think a lot of what we want in the world requires actually running experiments in the world.

So you'll hear things like, AI could be so great at identifying molecules for drug development.

And so maybe it could, right?

Maybe it would be much better than we are at identifying molecules to test.

But then you still need to run all these phase three trials, and one and two trials for that matter and animal trials and everything else.

And I think something we know from that area is a lot of things we think might work out, don't work out.

So I find it quite untested this question of will AI be this huge scientific accelerant?

There are things where prediction, like protein folding, could be a really big deal.

And it's also possible that just a lot more needs to be done of running experiments in the real world to make fundamental breakthroughs in things that would change our reality.

So the scientific side of this, I consider plausible and exciting.

And I also find it a little bit hand-wavy.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

I think that the place where this is going to have really rapid effects, because let's think about what these systems really are right now.

They are these large language models that are unbelievably good at impersonating humans and giving you predictive answers based on the huge corpus of human text.

And the problem with them is that they know how to predict what the entire human internet might say to something, but they don't really know if what they're saying is true or not.

I mean, even the word know there is a really weird word.

They don't know anything at all on some level.

And so I think you have to ask what would really work for a system like that, where it can be really brilliant, but it hallucinates a lot.

And I think the answers there have to do with social dimensions.

We have a lot of really lonely people in society.

I mean, loneliness is a true and profound scourge.

And I think what you're going to get from AI, before you get things that are economically that potent or scientifically that potent is a lot of companionship for better or for worse.

And this is, of course, a complete mainstay of both science fiction and fantasy, right?

Robot friends in science fiction, you know, you're running around to see 3PO and R2-D2.

You have daemons and other kinds of familiar beasts in fantasy.

I mean, when I was growing up, I was obsessed with this fantasy series called The Dragon Riders of Perne and I was a lonely, bullied kid and Ruth the White Dragon and the relationship between Ruth and Ruth's writer was like really important to me.

And we have a lot of lonely older people.

We have a lot of lonely young people.

And we also just have a lot of people who would like more companionship, more people to talk to.

I mean, again, the movie, Her, is a remarkable dramatization of this.

I could imagine ways that get stark, right, if people begin preferring AI relationships to human relationships, that could be a problem, but it could also not go that way.

One thing I found moving, there was a good piece in New York magazine about Replica, which is this company making, you know, what are at this point, quite rudimentary AI companions and you know, there are a lot of people named in the piece saying, I prefer this companion to people in my life.

But there are also a lot of people who said having this companion has given me more confidence and has given me more encouragement and incentive to go out and have experiences myself, right?

I'm going to go out and learn how to dance.

I'm doing this hobby because I have this supportive figure, right?

And I think a lot of us know this in our own relationships.

When you have supportive people in your life, it is a base from which you venture out into the world.

It isn't something where it's like, okay, I got two friends and a partner.

I'm never talking to anybody ever again.

And so I think there could be a lot of value in companions, and I think the systems we're

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

building in the short term look more like that to me.

The fact that a companion might say something that isn't true, I mean, my friends say untrue things all the time.

That is a little bit different than, you know, you're not going to have an AI doctor who occasionally just hallucinates things at you, right?

The liability on that alone would be a nightmare.

So that's a place where I think there's some real value.

I do think creativity and just an expansion of human capability is a real value, right?

The idea that I can't code, but I can create a video game.

I can't film, but I can make a movie.

I mean, that's cool.

We might be able to see really remarkable new kinds of art in it.

And then I think you're going to have a world for quite a while, which is just you have a team of assistants at no cost.

So I mean, right now, if you have a lot of money or you're high up in a firm, you know, maybe they'll hire you a chief of staff, an executive assistant, maybe you have people who you can outsource your ideas to, and, you know, they'll come back with a presentation and then you can give them feedback.

All of a sudden, everybody's going to have a team, maybe not everybody, but a lot of people have access to functionally a team, people who can research things for you.

You know, I want to think about how I could be healthier in this way, but I don't have a lot of medical literacy.

Like, can you look into this thing for me?

That's hard to do right now.

It's not going to be hard to do for very long.

And so, you know, I think if you just think about the way the economy works, right, you know, when you always have that line, the future is already here, it's just unevenly distributed.

Which powerful people already have large teams of people who help them live better?

Much of what those people do is remote at this point.

And when I say remote, I just mean it can be done on a computer, right?

Maybe the person is actually in the office, but you're telling people to go do intellectual work for you.

In a world where everybody has access to a lot of AIs like that, I mean, that might be quite cool.

And amazing things could be unlocked from that.

Are you familiar with the parable of the boiled frog?

I'm Estette Herndon, the host of the run-up from the New York Times.

I believe so, but I would need to remind you.

Give it to us.

For six years, I've had a hard time getting establishment Republicans to say publicly what they've all said privately about Donald Trump.

Then the midterms happened.

So you can put a frog in boiling water and it'll jump out, right?

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

But if you put a frog in a pot of water and then turn up the heat, he'll eventually cook to death without jumping out, right?

So we became that frog.

For the past few months, we've been reporting from inside the political establishment as they try to figure out their path forward behind closed doors.

This season, we go behind those doors.

You can find new episodes of the run-up on Thursdays, wherever you get your podcasts.

I want to talk about the other side of some of this utopianism, though, and even some of these middle ground scenarios, because I would say probably the most common kind of email we've gotten over the past few weeks is people really concerned about how AI is going to impact the labor market, and specifically the kinds of knowledge work jobs that tend to be done by folks with college degrees.

So you were mentioned things like art and research and copy editing.

As things that we could have these systems make a lot easier for us, but there are also a lot of people doing those jobs right now.

We've had lots of copy editors, writers, artists, programmers, emailing in wondering if they'll still have jobs.

We've had students like KDW asking whether it still makes sense to get a law or master's degree when the future of the economy is so uncertain.

And there's really two levels to this I've seen.

One is financially, am I going to have a job?

Will I be economically okay?

But then also on a deeper, more existential level, I think there's a lot of concern.

And I feel this too about what would it mean for me and for my life, my sense of self-worth, my purpose, my sense of meaning to have AI systems be able to do my job better than I can.

And so I'm wondering how you think about both dimensions of that, like how these systems could affect the labor market and how people should think about their effect on the labor market.

But then also this deeper existential question is raised of what it means to be human in a world where machines can do a lot of what we define as being human better than we can.

Yeah, those are profound questions.

And yeah, I've seen a lot of what I would describe as AI despair in the inbox.

I have a lot of uncertainty here as I do in everything I'm saying.

But I tend to be much less confident that AI is going to replace a lot of jobs in the near term than other people seem to be.

In part because I don't think the question is whether a job can be replaced.

There's also a question of whether we let it be replaced.

So this will be an analogy that will make some people mad, specifically doctors.

But there is a lot that doctors currently do that can be done perfectly well by nurses and nurse practitioners and physician assistants.

But we have created regulatory structures that make that really hard.

I mean, there are places where it's incredibly hard just to become a hair cutter because

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

of the amount of occupational licensing you need to go through.

We do not just let in many, many, many, many cases jobs get done by anyone.

We do let some of them get outsourced and we've done that in obviously a lot of cases.

But again, think about telehealth and how many strictures are on that.

Now we're seeing a little bit more of it.

So I am skeptical that AI is going to diffuse through the economy in a way that leads to a lot of replacement as quickly as people think is likely.

Both because I don't think the systems are going to prove to be for a while as good as they need to be for that.

It's actually very, very hard to catch hallucinations in these systems.

And I think the liability problems of that are going to be a very big deal.

Driverless cars are a good example here where there's a lot they can do.

But driverless cars are not going to need to be as safe as human drivers to be put onto the road en masse.

They're going to have to be far, far, far safer.

We are going to be, and I think we already are, less tolerant of a driverless car getting in a crash that kills a person than we are of human beings getting in a crash that kills a person.

And you could say from a consequentialist perspective or utilitarian perspective, maybe that's stupid.

But that is where we are.

We see that already.

And it's a reason driverless cars are now seeming very far off still.

We can have cars.

I mean, all around San Francisco, you have these little Waymo cars with their little hats running around, but they are not going to take over the roads anytime soon because they need to be not 80% reliable, not 90% reliable, not 95% reliable, but like 99.99999% reliable.

And these models, they're not.

And so that's going to be true for a lot of things.

We're not going to let it be a doctor.

We might let it assist a doctor, but not if we don't think the doctor knows how to catch a system when it's getting something wrong.

And as of now, I don't see a path given how these are being trained to not getting enough wrong that we are going to be fully comfortable with them in high stakes and even frankly, a lot of low stakes professions.

Now do I worry about things like copywriting?

I do.

I don't know how I think that's going to look.

And it's also possible it's going to make copywriters much more efficient and cheaper.

And as such, dramatically increase the market for copywriters.

I mean, the canonical point here is that we have more bank tellers than we did before ATMs because ATMs made it possible to expand banking quite a bit.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

And as such, even though you need bank tellers for fewer things, it did not lead to bank tellers being wiped out in the way people thought it would.

So these things often move through society in ways you don't really expect.

They create new markets.

They create new possibilities.

If they make us more productive, that creates more money.

But I get it at the same time.

One of the things that both worries and interests me that I think these systems are forcing or going to force a reckoning with, and this goes to your second point, Richie, how do

I say this?

There's been a strain of commentary and pushback from people saying that as we think about AI, we are dehumanizing ourselves in order to adapt ourselves to our own metaphors.

There's a point in a way that Megan O'Geeblon makes in her truly fantastic book, *God, Human, Animal Machine*, that metaphors are bi-directional.

You start applying a metaphor to something else, and soon enough, it sort of loops around and you're applying it to yourself.

You have a computer, and the metaphor is like the computer is like a mind, then you begin thinking your mind is like a computer because you get so used to talking about it that way.

And so you'll see these things, Emily Bender, the linguist, has really pushed on this.

You can see a YouTube presentation of her on AI and dehumanization, and she has a lot

of points she's making in that, but one of them is, people will say, and Sam Altman, the head of OpenEye said, to paraphrase, we're all stochastic parrots, with the point being that there's this idea that these models are stochastic parrots.

They kind of parrot back what human beings would say with no understanding, and so then people turn and say, maybe that's all we're doing, too.

Do we really understand how our thinking works, how our consciousness works?

These are token-generating machines, they just generate the next token in a sequence, a word, an image, whatever.

We're token-generating machines.

How did I just come up with that next word?

I didn't think about it consciously.

Something generated the token.

And a lot of people who do philosophy and linguistics and other sort of related areas are tearing their hair out over this, right, that in order to think about AI as something more like an intelligence, you've stopped thinking about yourself as a thicker kind of intelligence.

You've completely devalued the entirety of your own internal experience.

You have made value lists, so much of what happens in the way you move through the world.

But I would turn this a little bit around, and this has been on my mind a lot recently.

I think the kernel of profound truth in the AI dehumanization discourse is that we do dehumanize ourselves, and not just in metaphors around AI, we dehumanize ourselves all the time.

We make human beings act as machines all the time.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

We tell people a job is creative because we need them to do it.

We need them to find meaning in it.

But in fact, it isn't.

Or we tell them there's meaning in it, but the meaning is that we pay them.

So, you know, this I think is more intuitive when we think about a lot of manufacturing jobs that got automated, you know, where somebody was working as part of the assembly line and trying to put a machine on the assembly line, and then you didn't need the person.

And that is actually true for a lot of what we call knowledge work.

A lot of it is rules-based, a lot of the young lawyers creating documents and so on.

We tell stories about it, but it is not the highest good of a human being to be sitting around doing that stuff.

And it has taken a tremendous amount of cultural pressure from capitalism and other forces from religion to get people to be comfortable with that lot in life.

You have, you know, however many precious years on the spinning blue orb, and you're going to spend it writing marketing copy.

And I'm not saying there's anything wrong with marketing copy.

I've written tons of marketing copy in my time.

But you got to think about how much has gone in to making people comfortable or at least accept that lot.

We dehumanize people, and I wonder, you know, and I don't think this in the two or five or 10-year time frame, but on the 25, 50, 100, 150-year time frame, if there's not a possibility for a rehumanization here, for us to begin to value again things that we don't even try to value, right?

And certainly don't try to organize life around.

If I tell you that my work in life is I went to law school and now I write contracts for firms trying to take over other firms, well, if I make a bunch of money, like, great work, you really made it, man.

If that lot of great came from a good school and, you know, you're getting paid and you're getting that big bonus and you're working those 80-hour weeks, fantastic job, you made it.

Your parents must be so proud.

If I tell you that I spend a lot of time at the park, I don't do much in terms of the economy, but I spend a lot of time at the park, I have a wonderful community of friends, I spend a lot of time with them, it's like, well, yeah, but when are you going to do something with your life?

Why are you just reading these random books all the time in coffee shops?

I think that eventually, from a certain vantage point, the values of our current society are going to look incredibly sick.

And at some points in my thinking on all this, I do wonder if AI won't be part of the set of technological and cultural shocks that leads to that kind of reassessment.

Now, that doesn't work if we immiserate anybody whose job eventually does get automated away. Right?

If to have your job as a contract lawyer or a copy editor or a marketer or a journalist

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

automated away is to become useless in the eyes of society, then, yeah, that's not going to be a reassessment of values.

That's going to be a punishment we inflict on people, so the owners of AI capital can make more money.

But that is a choice.

It doesn't need to go that way.

Lots of people, Darren Acemoglu and Simon Johnson, the economists, have a new book coming out, *Power and Progress*, on this point exactly.

It doesn't need to go that way.

That is a choice.

And I think this is a quite good time for more radical politics to think about more radical political ideas.

What I hear you saying is that a huge question for all of us is not just the question, like the economic questions around labor markets, but the cultural questions about what we as a society choose to value and what we value people for.

And I totally on board with basically everything you were just saying, but I also think a lot about, for example, the now famous Keynes essay, *Economic Possibilities for a Grand Child*, where he was making a prediction almost a hundred years ago that around the time of our lifetimes, we would have reached a level of economic productivity that could allow us to work 15-hour weeks and that we were approaching this post-work utopia.

And we hit the productivity numbers and we're working not as much as we were 50, 60 years ago, but still a lot and a lot of the people who are most educated are working a lot.

So I guess I'm just wondering, why do you think that didn't happen?

What do you think it would take to actually culturally shift us in that direction?

And are you actually hopeful about that possibility?

Well, one thing I think it's commonly believed Keynes got wrong in that essay was he was interested in the question of material sufficiency.

What would it mean if you held material once steady, but increased productivity and income by this much, then how much would you have to work?

But it turns out you don't hold material once steady.

You have huge amounts of above inflation, cost increases in things like healthcare and housing and education, but also people want bigger homes.

They want to travel.

They want nice cars.

They want to compete with each other.

A lot of spending is also positional, probably doesn't make us happy, but there's an old line that the question of whether a man is rich relies on how much money his brother-in-law makes, which I think gets it something important.

So one version of this is to say, well, if you believe that AIs will create so much material and economic abundance that it just makes that kind of competition ridiculous, then people compete on other grounds that we're not going to get away from at least a somewhat competitive society.

People still on power.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

They still want to be partnered with and attractive to the people they want to be attractive to. Maybe everybody's going to spend 47 hours a week in the gym or something.

But I think the bigger question here is that the case gets something's wrong, but it gets others right.

And we know over and over again that humanity actually does go through very profound shifts in what it values.

Now, it doesn't do it in any given 90-year time frame, but it does do it in terms of the shift from monarchies to more kinds of democracies and more kinds of political systems did it in the shift from hunter gathering to monarchies and cities did it in the shift to agriculture.

Religions create a lot of this.

I mean, I think just like I have no predictions here, but I think that the question of how religions both old and new interact with dramatic changes here in the world is going to be very, very, very interesting.

And I think a lot of them have a lot to say about these questions of how we value human life that is simply waiting there to be picked up.

I did this episode on Shabbat not long ago about Shabbat and rest, and the idea that a day of rest is the day that should be the way the rest of the world works and that the Shabbat practice in its radicalism is a profound critique of the values of our economy as they exist right now.

I think you can imagine that becoming much more widespread, that becoming a much more profound practice and cultural, not just artifact, but challenge in the kind of world I'm describing. So I don't believe in utopias just in general, but I do believe in change.

Now it's not going to happen.

I don't believe typically change happens so quickly that between when I am 38 as I am now and when I am 50 or 55, that will have stopped having this overwhelming ideology of productivism, nor that I will have stopped applying it to myself.

I mean, I have completely imbibed the values of this culture and standing outside them to critique them in a podcast is a lot easier than not weaving them through my own soul.

But I don't think the fact that Keynes was wrong about how much we would work and what we would want means that these kinds of shifts don't happen.

I think that a longer view should make that look pretty different to us.

And you never know when you're on the cusp of a world working quite differently than it has in the past.

For

all we've been talking about how AI could change the nature of the economy, in our conversations you've been a lot more skeptical or at least hesitant about whether sort of on net AI will lead to the kind of takeoff in economic growth and productivity that a lot of people think it could.

You mentioned earlier that you're skeptical of whether AI will lead to a sort of super takeoff in scientific progress.

But there are lots of other ways you can imagine AI systems making us a lot more productive a lot more efficient.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

We've already discussed things like automating a lot of repetitive work.

So could you just unpack why you're a bit more skeptical about whether AI will supercharge economic productivity?

So one thing I would say is one of the possibilities I hold is that they won't as I keep trying to emphasize I'm open to a lot of things being potentially true here.

But yeah, let me give two reasons.

If I was trying to imagine 15 or 20 years from now when people are like to paraphrase an old line about the internet, how can you see the AI revolution everywhere but in the productivity numbers?

Why am I think that would be?

So one reason is that systems that don't really have an understanding of what they're telling you that they have this capacity to predict the next token in a sequence that's going to turn out that there is an ineradicable amount of falsehood and hallucination and weirdness in there that just makes it too hard to integrate this into the parts of the economy that really make a lot of money.

That you would need such a level of expert oversight of them, somebody who knew everything the system knew or needed to know so it could know when the system was telling it something untrue that it just doesn't really net out.

So that's one.

But the one I think is maybe even more likely, think about the internet.

Let's say that we go back in time to 1990 and I tell you what the internet is going to be in 2020, the size of it, the pervasiveness of it, the everywhere-ness of it, you will have in your pocket, in your pocket, imagine a computer with functionally the entire corpus of human knowledge on it.

You'll be able to search that in a second.

It will talk to space and you'll be able to talk to and collaborate with anybody anywhere in the world instantly.

You will pull this all-knowing pocket rectangle out, press two buttons and the face of your collaborator in Japan will appear immediately and you can translate.

So in addition, you can now work with anybody.

You can read anything in any language.

And I said to you, if we had that technology, what do you think would happen to the economy?

Think about the amount of knowledge it is now instantly accessible.

Think about the amount of collaboration and cooperation that is now being unlocked.

Think about the speed.

I mean, you imagine a journalist going before to the library and going to look stuff up and now you can just Google everything.

What do you think will have happened to the pace of scientific progress?

What do you think will have happened to the pace of productivity growth?

And if you had given me that in 1990, I mean, I would have been sick.

So I probably wouldn't have had a very good answer, but I think if you had framed that in 1990, somebody would reasonably say, wow, that is going to hypercharge the economy.

That is going to hypercharge scientific progress.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

And here we are.

And it did none of those things.

The productivity growth has been quite disappointing in the age of the internet, worse than before it, you know, in the sort of post-World War II period, you know, there's a lot of people and we've had some of them on the podcast worried about the slowdown in scientific knowledge. The advances we are making seem less potent in many ways than the advances we made before. And obviously there are a million different explanations for this, but one explanation I favor more than other people seem to is that we weren't really wrong about what the internet would do to make us more productive.

There is no doubt in my mind that I am profoundly more productive than I could have been before it.

What we were wrong about is the shadow side of the internet, is what it would take away from our productivity.

So now, like, go back to that 1990 thought experiment and let's say I come to you and I say, we're going to invent a technology.

Everybody's going to have it on them at all times, all times.

And what it's going to do is it is going to have so much content, so much entertainment and so much fundamentally distraction that all of humanity averaged out.

They're going to be 30 to 45 percent more distracted than they are now.

They're going to be angrier, they're going to be more annoyed, they're going to be more tired, they're going to be less able to hold a train of thought.

The time they can spend focusing their attention on one thing is going to reduce.

The amount of time for reflection and contemplation is going to narrow.

What do you think that will do to the economy?

Or scientific progress?

I think somebody would properly say, oh, that seems bad.

And both those things in my view happened.

There's been a productivity enhancing effect of the internet.

I just don't think you can do a job that is online and not see that.

And there's been a productivity reducing effect of the internet.

The number of times I distract myself from a train of thought in an hour, in one hour, by flicking over to some garbage in my email or looking at Slack, because you've told a joke, you don't really say much on Slack actually, Richie, it's constant.

And I think that has a real cost in the depth of what I'm able to produce.

You're welcome for not imposing those costs on you.

I appreciate that.

And I think that it's really possible that happens with AI.

And this was part of the Gary Marcus interview we did a couple months back.

I think it is really possible that particularly large language models are going to prove to be a much better source of distraction creation than of actual productivity enhancement.

Because particularly if we don't quickly get to the kind of AI that can make profound scientific advances on its own, that is truly autonomous and generative in that way, what I think we're going to get instead is AI that is, as social media, which has a lot of machine learning

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

beneath it, is really, really good at finding and now creating content personalized to us to distract us, serving us up exactly what we want.

I think the bigger and more plausible set of distractions are actually social companions. Look, I sometimes go to a coffee shop and I work there with my best friend.

When I do that, I really enjoy the experience and I'm less productive because I enjoy talking to my best friend at the coffee shop.

And if you have your AI best friend and also your AI girlfriend or boyfriend or non-binary sexual partner in your pocket at all times, how distracting might that be?

What will it mean to have a lot of those kinds of figures in your life or entities in your life?

So I think it's entirely possible that this is a technology that for good or for bad ends up being a technology of entertainment, of socializing, of content creation, of distraction, much more than a technology of productivity that particularly if it is still human beings who have to be the ones who are productive.

So far we've seen a couple of these studies come out and they're things like, we gave a bunch of coders GPT-3 or GPT-4 and it turned out they're a 35% more productive.

I don't buy those studies at all.

I don't buy them at all because if you put people in sort of like a laboratory condition and you just tell them to use the new technology in the most productive way, yeah, it's going to make them more productive.

But that's not what's going to happen necessarily when these same people live in the world.

We're on the one hand, yes, you could use GPT-4 or 5 or whatever to help you code for long stretches or you could use it to screw around on the internet and play video games that are so immersive and personalized and we've never seen anything like them before or talk to this perfect just for you AI companion who's now like always there in your ear.

I'm not sure that I think people are ultimately going to be that much more productive.

So I think there's a lot of ways this could go weird but also that it might not be the economic boon we are hoping it will be.

I will just say that Ezra is constantly distracting me on Slack and that I actually think a companion chat bot might be an improvement so I'm not sure I'm a little skeptical, it's reasonable.

But I will say, in the background of a lot of these questions, including the one you just answered in I think a pretty deep way, is also this question of and what you were just saying makes me think maybe there isn't a clear answer, but there's this question it keeps bringing up for me of like what kind of technology do you think AI is?

You mentioned earlier that you don't like the social media analogy and you think it could lead to regulation that maybe is inadequate and I'm wondering if you think there are analogies that are better because there are a lot of people who are making these comparisons right now.

Some who say like AI will be the new internet, others who compare it to sort of these general purpose technologies like electricity or the steam engine which sort of became like the foundation of economies, you'll hear it be compared to the printing press, to oil, to fire.

And clearly, these analogies are important because they sort of determine how we treat

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

the technology and how we try to regulate it.

So I'm wondering if there is any good analogy for like what this technology is and what it could do and if so what that analogy is.

It's funny, I was actually just talking about this with somebody yesterday, Henry Ferrell who's a political scientist and has done really, really cool work on everything but AI among them.

He has a good piece on high tech modernism, people should look up.

And we were talking about this question of analogy and what is the analogy and the one I was saying was what if we fully understood the social implications of simultaneously the internet and the globalized opening of China particularly but you could say sort of other big manufacturing exporters as of 1990.

What if we could feel all that in its full weight right then?

We knew what was coming, what would that have meant, what kind of preparatory work would we have wanted to do that this feels to me like it is on that level, right, like internet plus China.

Now maybe that's wrong and in the long run maybe that's too small but I think something like that, I don't see it as a technology.

I see it as a, in the sort of Timothy Morton conception, a hyper object, a thing that kind of touches everything in weird ways such that it also changes the way you look at the things it's touching to our whole earlier conversation about the metaphors on the human mind or social relationships.

So the internet is helpful to me there because the internet had that quality of touching almost everything.

And I think China is helpful there because China really changed the way the US economy worked.

I think much more frankly than the internet did.

And I'm using China again a bit as a stand in for globalized export based supply chains but China was like the big mover there.

So maybe like that but I don't think anything in particular is going to work for something that will be as weird as this is going to be.

Like if it is true that we are creating inorganic intelligences that are going to surpass our generalizable intelligence within the lifespan of a human being born today, we're entering something profoundly new then.

Now maybe it's just weird, right?

I mean there's a lot of sci-fi and I think the Star Wars universe is a good easy at hand version of this.

We're just like the world is full of computers, you know, computer companions and they talk and they beep at you and they're like friends and they do things that are useful and everybody just goes along their day as normal and like maybe that's where we're going.

I mean there's been a lot of imagining of that.

It really does depend whether or not we get things that are super intelligence level or things that are kind of human but programmed.

And I don't know, I think again, uncertainty is a place to sit right now.

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

So I think that's a good place to start to come to an end and give everyone a break from all the AI talk.

So let's do some recommendations.

I know you just got back from a music festival.

I was wondering if you have any music recs for the audience.

I would love to do some music recommendations, what I actually enjoy.

So I saw a show at this festival, in fact, by Daniel Ponder, who is this just, I mean, to have a voice like that, I don't know.

Her album is called Some of Us Are Brave and a song maybe to start with there is So Long.

But it shows the kind of nerd I am that I was sitting there also thinking about AI and thinking about all that we miss in the importance of human beings being in a room together and actually experiencing the beauty they can create while I was listening to her.

Because I mean, I can totally imagine a world where we're generating a lot of computer art, but it is not going to have the effect of hearing about her life from her and then hearing her sing, her art.

I mean, that was really, that was really quite special.

Yeah.

So I've been listening to a lot of Felix Rosch, R-O-S-C-H.

He's a modernist, composer, electronic, but a lot of strings, actually a lot of music

I like has that quality, electronic with a lot of strings, like lightly electronic with a lot of strings.

But his work is really, really beautiful.

The song that got me into him was this song called In Memory of a Honey Bee, which I really, really, really recommend checking out.

But his song Clouds is great, Driven is great.

Just go check out any of his essential mixes.

And I found this artist recently, Mabe Frotte, and I might be saying that wrong, but you spell it M-A-B-E-F-R-A-T-T-I.

And cellist and electronic elements and just some of her songs are just astonishingly, just truly astonishingly beautiful.

And then going back, I guess, to the Rick Rubin episode, he ended up in that episode recommending a Nils Fromm album.

And one I ended up listening, spending a lot of time with before that episode and after is the one that he did with Oliver Arnold's called Trans Friends, which is a fully improvised album in one night.

But I just love that album and would urge people to give it a listen, just beautifully meditative, deep, rich music that also has a very different feel to it, knowing that it's two friends who decide to spend a night together with a bunch of their instruments, and this is what they came up with.

But I appreciate it for what it gave to the world.

Well, given that description, I will definitely be happy to check that album out.

Last question before I let you go, is there anything that's on your mind right now as you think about your life and the show over the next year that you'd want to share with

[Transcript] The Ezra Klein Show / Why A.I. Might Not Take Your Job or Supercharge the Economy

the audience or talk about?

Ooh, do I want to share it with the audience?

I'm about to move from California to New York.

And I think if you listen to the show, you can get a sense that my identity is a Californian one's really deep and that there's a lot I draw strength from here and curiosity from here and a lot in my intellectual way and personal way of being in the world that fits here.

New York has never been my place in part because I have spent most of the time I've been there in midtown, which is certainly not, I think, the finest part of it.

But so it's going to be a big move and I'm trying to remain open to what the place has to offer.

Obviously, other people really love New York and open to recommendations about great things in New York.

But there will be some, probably disruption in the show as I make that move.

And also just some, hopefully, interesting changes to me as I try to absorb a new place with a different culture.

I will just say as a fellow native Californian, I think you're making a terrible decision, but I am excited for you and for what this next chapter brings.

So thanks for having me.

This is a lot of fun and I hope we get to do it again sometime with you on the East Coast.

Thank you.

All right, this episode of the other clown show was produced by Roger Karma, Kristen Lynn and Jeff Geld.

It was hosted by Roger Karma.

Thank you to him.

Fact-checking by Michelle Harris and Kate Sinclair, mixing by Jeff Geld, original music by Isaac Jones, audience strategy by Shannon Busta, the executive producer of New York Times, depending on audio as Andy Rose Strasser and special thanks to Sonia Herrero and Christina Samuelski.