OpenTheo

Human Identity and the Meaning of Artificial Intelligence | Joshua Swamidass

March 2, 2021

The Veritas Forum

A Conversation with a Secular Humanist, Jurist and Law Professor, Ryan Calo, and a Christian, Scientist and Bioengineer Professor, Josh Swamidass, about the philosophical and biological implications of artificial intelligence. Presented by the Veritas Forum at University of Washington. • Please like, share, subscribe to, and review this podcast.

Transcript

Welcome to the Veritas Forum. This is the Veritas Forum Podcast. A place where ideas and beliefs converge.

What I'm really going to be watching is, which one to have the resources in their world view to be tolerant, respectful and humble toward the people they disagree with. How do we know whether the lives that we're living are meaningful? If energy, light, gravity and consciousness are in the street, don't be surprised if you're going to get an element of this in God. Today we hear from Washington University and St. Louis Professor of Pathology, Immunology and Biomedical Engineering, Joshua Swamadas, as well as Juris and Law Professor at the University of Washington in Seattle, Ryan Caleb, as they discuss human identity and the meaning of artificial intelligence.

Moderated by Rebecca Rice, Associate Professor and Chair of the Department of Philosophy at Seattle Pacific University, presented by the Veritas Forum at the University of Washington. Thank you for having me. My name is Joshua Swamadas, as you heard.

I'm going to talk to you about the meaning of artificial intelligence and how we think about it in terms of human identity and what it means to be human. My goal really here is to throw some things into the room that will hopefully create some interesting conversation and show you how I've been thinking about these things as a scientist that works with artificial intelligence for a while. It's interesting because when I first started my career, I graduated undergrad in 2000.

That was a while ago, right? So I guess I'm in a different generation now. Is that true? I hope you don't think I'm too old, do you? Back when I was in your age, machine learning was actually a thing, and I knew I wanted to do it, which is why I went to grad school for that. But it wasn't anything like what it is now.

And in fact, a lot of the things that I'm going to share with you now, I didn't even know that this is the way things were going to be. But there were a couple of grand questions that were visible way back then, even when I was in school, that are still with us today. And those questions are still unanswered.

And in fact, I think they're probably going to be still unanswered 20 years from now. And so I don't really actually have a set of answers to give you, but what I want to tell you is what some of those questions are and some of the interesting things that we found along the way, and maybe even hear what you think about them. So I'm going to need to be three unanswered questions.

Is it possible for us to build machines with consciousness? Is it possible to build an artificial mind? No one knows the answer to that question. No, you see it in movies all the time, but no one really knows the answer. Here's the second question.

If we could, how would we do it? How would we build that truly conscious machine with an artificial mind? Even the people who think that we can can't tell you how to do it yet. Isn't that interesting? And it's tricky because maybe we'll never know until we do it, but how do we do it so we don't know how to answer that basic question. And this is probably one of the more fundamental things, is how do we know if we succeeded in building a conscious mind? No one knows how to even answer that question.

These are three pretty big questions, right? And I'm framing it in terms of machines because artificial intelligence is really about how machines are thinking and doing things more and more than we ever thought possible. And I'll even say that even though I don't know the answer to these, we're far more likely in your lifetimes to encounter an intelligent machine, like a machine with a mind than we are to encounter an intelligent alien. You agree with me on that at least? So it seems like that might be worth something like, okay, maybe not.

Maybe we'll encounter an intelligent alien so far away we could never actually meet them face to face, but the chance you'll stand face to face with a machine that has a mind is higher than that. You're going to be that at least, right? Okay. So let me kind of explain to you some of the key people and some of the things that have happened that people are making people wonder about this question.

So Alan Turing is considered like the father of modern computers. There's a really good movie about him, which I really recommend, but he was really thinking about this. And actually people were thinking for a long time about, you know, if you could actually make

a universal machine, which is called the Turing machine after him, could, is consciousness computable.

And Turing thought that it was, because essentially he thought that we were basically just thinking machines, and so we could make a machine that could do any sort of computation, then we could make a thinking machine. And as crazy as that sound, he's the guy who actually had the vision to think about something like your cell phone or computer that actually is a single piece of hardware that just by downloading new types of programs on it, it essentially is a universal, like, switch, you know, like a Swiss army knife. It can do anything in a certain sense, right? Anything that is computable, it can do.

So how do we solve this problem? Well, Turing took a very, like, a pragmatic approach. He tried to make this as scientific as possible. He said, figured that if we can basically put, make three rooms where there's a person, A, and there's a person acts in one room and a machine acts, and then you're basically allowed to chat.

So he's envisioning Twitter, or AOL instant messenger, but that dates me too much, but you get the idea. Google chat. If you can Google chat with a machine, you can Google chat with a person, and if the person who's trying to determine it can't determine which one is which, then we know that a machine is conscious.

Okay, that's his idea. But already you should be a bit skeptical of this, because we already have machines that can sometimes pull this off. There's a really interesting example of actually just recently how Google assistant, like the technology behind Google assistant, that all our phones are sometimes linked into.

It could actually fool a person at a restaurant by just acting and just having a conversation to make a reservation. And the person on the other side actually thought that it was an actual human. It was actually just, you know, an AI machine.

And no one I think thinks that Google assistant has a sense of human awareness in a mind. Do you think that? And it can start to pass that. And that's the fundamental problem, because even if you can emulate all the actions of a mind, does that mean you actually have a mind? And what really gets to is this fundamental problem is that we all know that at least we ourselves have minds, because we have direct firsthand experience that we have a mind, right? You perceive that directly.

And when it comes to actually perceiving other people's minds, that's a different question. But we can always make this inference that plan ago would call proper and basic that other people have minds too. But what do you do with a machine when you don't actually know, by the way? There's not a good way.

John Searle is a philosopher who really formulated that problem in what he calls the Chinese Room experiment, a thought experiment. So imagine a person who's following a

rule out of a rule book to interpret squiggles. He doesn't know anything about Chinese.

This is a very Western story. So if you actually understand Chinese, this will make no sense to you. You can kind of flip it around.

Imagine a dumb American in the room. Is that hard? [laughter] Imagine a dumb American in the room who doesn't know a thing about Chinese, or what this is. He just see squiggles on a page, and he's looking at a rule book and just manipulating these symbols.

And then there's Chinese people who are kind of putting stuff into this room. He's manipulating symbols, or maybe there's a group of people doing it. And then they're spitting out stuff.

But through this process, it actually seems like he can have a conversation. And then Searle asks the question, "Well, does this man actually understand Chinese?" Right? And that's a fundamental problem, because most of us would say, "No, he doesn't understand Chinese." And certainly the book he's reading stuff out of doesn't actually understand anything. It might encode information that he's manipulating.

So he's really just succeeded in fooling the people outside that he understands Chinese when he doesn't. And that's the fundamental problem with the Turing test. So you can imagine a case where a machine can fool someone to think he has a mind when it really doesn't.

Now there's a couple other advances that have really pressed the challenge of this problem. One of them is advances in neuroscience. So, everyone since the beginning of history has realized that humans have a mind.

Where that mind resides has actually been widely contemplated different places. Not everyone thought it was in our brains. Like Egyptians, they would actually be more concerned about keeping the body intact than it was about the brain.

So sometimes they would actually bury people without their heads. Because that wasn't the important part. Isn't that interesting? Anyways, we can't imagine that because we know that the brain is somehow connected to the mind.

About 100 years ago, a little bit more, actually more than 100 years ago, maybe about 130 years ago, Ramon Hikal, who was a Spanish scientist or anatomist. He found a way to actually do a silver stain so we could look at individual nerve cells underneath a microscope. And he's a guy who was considered the father of modern neuroscience.

He ends up getting a Nobel Prize. He was an atheist. He actually became a Christian later on in life.

He wrote this really beautiful book called Notes to Young Invest here, which is a little bit misogynistic, but everyone was a little bit at that time. So if you're a woman, I'm very sorry. But it's still worth real reading to just put asterisks over that part.

But he talks about what he does to advise what it means to be a scientist. But what he did is he actually looked and just was studying the shape of neurons in our brain and came up with the idea that maybe what's going on is these neurons are processing information. The dendrites are getting information, integrating it, and then the axon to send the out information.

That's how our brain works. And you know what? He's right. That's exactly correct.

And so as we start to understand more and more of this, here's the thing, though. The more and more we study neuroscience, the more and more it looks like the Chinese room. It looks like things that individually can't really give an account for a mind.

They're certainly connected to the mind somehow, but it's not enough to give an account for it. And I'll tell you what's going on with artificial intelligence is creating the same sort of challenge, too. So this is AlphaGo.

So when I was in college, it was the first time that Kasparov was beaten by Deep Blue, which was a computer that could actually play chess. A quick aside, when Alan Turing came up with the idea of computers and a guy first implemented, they were pretty certain that there's two problems they wanted to solve. They wanted to solve natural language to be able to add a computer to speak and think, or use human language.

And they wanted to get it to actually be able to play chess. And it's interesting what happened. They found out that getting it to understand English was not a ten percent of the time.

It was not a ten-year problem like they thought. We still haven't solved that problem. And they thought that teaching it to play chess would be extremely difficult.

You see what's going on? They're trying to make an analogy saying, well, everyone can speak English, so clearly that's easy to do. I mean, if you're tired, everyone can speak a human language, so that's easy to do. But chess, only a few people can do that.

And so that's harder. It's like a very human sort of way of doing that. But it turns out computers can solve chess really easily.

Now, they couldn't solve Go because a lot of the problems that were involved in Go had to do with pattern recognition on a higher level. You can actually just play it out in chess and look at tons and tons of examples. But that's where the thing called deep learning comes in, where it turns out that's really good at pattern recognition in a way that surprisingly maps in important ways, not identically.

And maps in important ways to how the human brain works. You can train a neural network in a computer science sense to be able to pick up fuzzy things that are just patterns that you can't fully articulate and understand. You can give like a computer an intuition.

It's encoded in numbers. But what is it actually going on? I mean, what I said are all very valid ways to describe it, but it's all just matrix multiplications. And simple math operations.

It's actually fairly disturbing how mindless the process is. I remember when I first thought about matrix, you guys learned about that? Like is it sophomore and high school or so? I remember thinking this is the most bizarre thing in the world. Why would you manipulate numbers in this way? And I said, no, it'll be very interesting and useful down the line.

It's very helpful for making sense of certain things in science. Hand-waving as high school teacher doesn't quite know how it connects. Okay.

The thing is that basic operation of a matrix multiply ends up being the fundamental thing behind the most powerful machine learning stuff of today. Isn't that crazy? And it just has to do with tuning and tweaking the numbers in these matrices that you can do the most amazing things. Like one of the questions I'd use in my research is trying to figure out if I can build computers that can understand chemistry.

I'm really interested in how to make drugs that are less toxic. I understand how the body metabolizes drugs or molecules when we take them into our system and how it changes them. And it turns out that we don't know the precise rules because it's not a simple system.

It has to do with interactions with proteins with small molecules and all of that. But what we can do is get a lot of data to find out how it has happened. And then it turns out that there's far too much data for a human to sit down and sensibly make sense of.

But I can build an artificial brain. I wouldn't say mine. I don't think it understands.

And maybe the way, I don't know. I just have consciousness. That's for sure.

But I can show it lots and lots of data. And then it can come back and actually do better than organic chemists. And I've been working this in a long time at understanding how drugs are going to metabolize and how they're going to become toxic.

I can give this insight. It becomes a tool to help us discover new things in science. That's one of the things that my group does.

These are just some schematic representations of what it looks like of how all these

matrix multiplies are happening. Every single line there is actually a matrix multiplied. There's other questions that really arise in this.

One question that arises is it possible that minds really just need to be embodied? There's a contrast between two science fiction stories here. One is her. Have you guys heard of that story? That's the one on the right.

It's just a computer program. It's like an operating system that's conscious. That's a disembodied artificial intelligence.

I would say the vast majority of artificial intelligence today is disembodied. What I was just talking about with the chemistry is disembodied. It's more like an operating system.

In this case, it has consciousness. But I think we have to ask, is that even a coherent idea? Maybe it's not. On this other hand, is that's Doris from Westworld.

Westworld is an HBO. You guys are probably too poor to afford it. It's a really good story.

What's going on here is that she looks like a human. She has a body. She has a brain that is actually silicon or whatever they made it with.

The story is actually over a couple of seasons of how they come to self-awareness. It's a deeply embodied process. They have multiple lives.

They keep on getting to have their minds wiped. They have memories that are given to them. There's also this memory that their body carries with them that ends up being really central to how this is done.

There's a lot of thought about it. Is it even really possible to make a truly intelligent machine if it doesn't have a sense of space and a sense of actually being embodied, for example, in a robot? Maybe it's not possible to have consciousness without having a body. There's another really great story which is a historically significant moment because the second season came out today for altered carbon.

Now you guys can afford Netflix, right? How many people here know about altered carbon? All right, more people know about that. This is three bodies that all have the same mind. It's Takeshi Kovacs.

In this story, basically, there's this technology stolen from angels that allows people to put little things in the back of their neck and pop from body to body to body. One of the core points, one of the few clear philosophical points in this incredibly philosophically interesting series is that immortality grants it to fallen people is an incredibly evil thing. What happens is there's this people called the myths, or yeah, myths, right? That bounce from body to body to body and they own everything and they take everything and you have to start wondering if they're even human anymore.

Even though they have a human body and a human mind are even really human, it's so far from a human experience that you have to wonder. In a way they aren't encased in a body, but it's not really embodied in the way we think about it. That's actually why I like artificial intelligence so much.

To be clear, there's no such thing as artificial mind yet. We don't know what's going to happen in the future. When we start thinking about these things, it really brings us to those grand questions.

Those grand questions that don't have an answer yet. You might think because they don't have an answer, we shouldn't think about them, but I think it's actually the other way around. We start thinking about these questions that brings us to this deep and powerful and meaningful dialogue that actually was around long before even artificial intelligence hit the scene.

You go back just a little bit of time, Darwin introduced Origin of Species almost exactly 160 years ago. The equal co-inventor of evolution, or co-discover, I should say, of evolution with him, was Alfred Russell Wallace. He liked the idea of evolution, but there was one point where Wallace, Wallace's doubt, just couldn't get around making sense of it.

It's one of the grand puzzles again. He couldn't give an account of how the human mind arose by evolution. It's still one of the grand puzzles to even figure out when it arose, how it arose, and it's certainly reasonable to wonder if there's more to it than that.

You go back farther, this ties into even my Christian. It ties into discussion about what exactly is the soul. These are grand mysteries.

It's part of what makes us human is not because we only care about questions because we can answer them, but because some of the most interesting and most important things to engage are the things that have been engaged by people long, long before us, and that actually probably don't have simple answers. That's okay. That's actually part of what it means to be human.

In fact, I think that is what the meaning of AI is. It brings us back to this question. What does it mean to be human? Thanks a lot.

[applause] Professor Palo. Thank you. That was fascinating, Josh.

Thank you so much. Thank you for the introduction, Rebecca. I'm really delighted to be here and really feel honored to be included in this forum given the great history of the organization, putting it on and the various sponsors.

But I want to be clear that I'm coming from this set of issues from a very different place. Not just that I'm coming at these issues as a person who does not identify as a Christian

per se, but also I'm a law professor. What I've always loved about the law and the study of the law, I was a philosophy major and undergraduate.

I just loved it. I thought very hard about getting a PhD in philosophy because I thought it was so interesting. One of the things that really attracted me to the law is that the law is a place where we come in order to set the actual rules.

There's this amazing article that I just loved so much by an author named Robert Covert. It's called Violence and the Word. It begins, legal interpretation takes place on a field of pain and death.

Pretty dramatic. What does he mean? He means that when judges make decisions, they are looking at deep questions about meaning. They are looking at questions of causation.

They are looking at philosophical quandaries. But then they have to give a decision. That decision has consequences because people's bodies and their money and their liberty are on the line.

The state has a monopoly on the lawful use of violence. To me, it's really interesting and important and actually something I take very seriously. The law is the place where we come together and we make decisions about when are we going to actually allow or disallow something to happen.

The consequences are quite material. What I want to say about reaction to Josh is really extremely helpful. What I want to say about it is that from my perspective in a sense whether or not machines can be like people and have a mind, it doesn't actually matter very much yet.

One day it could be that there will be some machine that we are unable to pass the Turing test with flying colors and we are unable to have any account for why this isn't just like a person. If and when that were to happen, it would be a deep challenge to the law because the law everywhere assumes a biological person. For example, imagine that you had this artificial intelligence that Josh built.

An artificial intelligence that a team built. It's Google's project after beating some video game and go. They make a person.

This machine comes forward and says, "Well gosh, I'm really glad to be here. I'm just like you and I'm really really smart." Everybody loves this machine and no one thinks of it as anything other than a real person. It's so popular that 20 other people that run for president of the United States.

We decide, "I think that we should elect this machine. It's going to do a better job than the person currently in office." Whoever that may be, it's time. The question becomes,

does that machine, if it was built in 2020, does it have to wait until 2055 in order to be the president of the United States? Because the Constitution says in Article 2 that you have to be 35 years old.

You see what I mean? So much about the law assumes a biological person. That there's a kind of incoherence almost. What I try to dramatize is there are so many interesting questions for the law that are well short of questions about what do we do about an artificial person for say.

If we just take our common experience, don't look at the law for a moment. The things that I experience at home is my wife did a ban on Amazon Echo, Alexa, but we had it for a short period of time in the house. The kids would learn to control Alexa.

They would think not only that they could just tell Alexa to shut up, but that it was kind of funny. Alexa, would you like to build a snowman? I have a six-year-old daughter and a frozen is extremely popular in my house. Then Jude and my son would come in and go shut up Alexa and they would giggle about it.

I thought to my wife and I would look at that situation and say, there's something a little uncomfortable about that. You don't really talk to people that way. We don't use those words.

You know what I mean? But then they just change the words that they used. It was this comfort because on the one hand, I don't like to think of my kids as being socialized to give orders and to not be respectful and to not use pleas and thank you. But at the other hand, I don't want them to believe that this machine, this machine made by a corporation, mostly to sell things should be treated like a person.

It creates this tension. It turns out that just as we might struggle to characterize something that presents us human even when it isn't, so does the law. So does the law.

There are these great cases. By the way, some of them have to do with a lot of them have to do with embodiment, but they don't all. But I'll tell you a story that one of my favorite stories is this.

In the 1990s, you're going to break my heart because not a lot of people. But how many people know what I'm talking about when I say chucky cheese? Is that something that... Okay, thank goodness. Sorry.

I don't think that went out of business like blockbuster. Yes. Good.

That was close. Okay. Because sometimes I got to explain to people like, "Once upon a time..." Why didn't I go out of business? I don't know.

Why didn't it go... Anyway, so chucky cheese when I was growing up was like a big deal.

I mean, you know, now that there's coronavirus, I'm never going to go to chucky cheese ever again. But the point of the matter is that it was a big deal for me and it was really interesting.

And the thing I loved most about chucky cheese was when all of a sudden, the animatronic robots would start up and they would like play a song. I mean, the new chucky cheeses don't actually have as many robots, but the old ones had a lot of... It was like a robot band and there was like a mouse and like a, you know, whatever. And I kind of slightly racist caricature of an Italian pizza guy who was like, "Hey, you know what I mean?" Like it was... Yeah.

And so... But the point of the matter is that this was like this group of robots and they would like play a song. Okay. Well, in Maryland in the 1990s, some really enterprising tax authorities caught wind of what chucky cheese was doing.

And they went to chucky cheese and they said, "Hey, looks to me like you're serving food during a performance. And therefore you have to pay a performance tax on food." And chucky cheese was like, "What? Are you guys serious?" And I'm like, "Yeah, yeah, there's a performance. There's a performance happening in your restaurant.

And there's a... in Maryland, there's a performance tax that you have to pay on food." You know? So they went to court. They went to court. Chucky cheese had to convince... a court had to decide whether or not robots could perform.

And it did this whole analysis of like, "What does it mean to perform? Does it require spontaneity?" Like some of the very things that Josh is talking about, this court had to think through. Not in as deep and interesting a way perhaps, but had to think through. And then what was the consequence? Well, the consequence was that chucky cheese would have to pay money or not.

Ultimately the court decided it's not a performance. It's not spontaneous. In fact, the court says with this sort of sweeping authority of a court, robots can't be spontaneous.

That's an intrinsically human thing, right? And it tell that to artists who over the years have built these very interesting, emergent, robotic things, like the robotic church in Brooklyn, and where no two performances are the same. And it's very much contingent on, it's very much spontaneous. But that's what the court said.

Another example... So do they want to pay the performance tax now? They ended up not paying the performance tax because... No, but for the robotic church. Should they have to pay the performance? They should have to pay the performance. They don't serve food.

But they should have paid the very lawyerly answer. But at any event, so this is a court case that they had to say. Another one I really like a lot is even earlier.

So in the 1950s we began to import these robot toys from Japan, right? And they're really famous. So there's these robot toys, these old robot toys, and someone will move around. And they had to figure out what the tariff would be.

What would the tax be on these robots? Okay? And so the struggle was that if you're import, for historical reasons, having to do with our relationship with Europe, if you're importing a doll into the United States, it was less money than if you're importing another kind of mechanical toy. You know what I mean? And the way that doll was defined in the law was that a doll represents something animate. And other mechanical toys don't.

And so the court, once again, had to decide, does a robot represent something animate? You know what I mean? And what is it to be animate? And they got into these deep questions again about what it is to be animate. I mean, really, you know, starting with the dictionary, but getting into like philosophy about what is it to be. And what the court decided, I had to do a little questionable, but what the court decided was that a robot does represent something animate because a robot is a mechanical man, a mechanical person.

However, a toy robot represents a robot. [laughter] And so it was charged with the higher tariff rate. What is my point with all this? My point with all this is that, you know, what the law cares about, and again, the law is going to be making the decisions about like, you know, who pays what and who goes to jail and what, you know what I mean? So what the law cares about is the quality of the object.

That it has a particular affordance. It causes a particular reaction in people. It doesn't care about whether the thing is alive or not alive or if meaning or not meaning, but it looks at, you know, so to simply, the law cares about whether an artificial intelligence is embodied.

Or not embodied. Because there is a different scheme for liability when you're physically hurt than when you just are hurt emotionally or hurt by loss business. So those are the inflection points.

And so what the law, so what it fascinates me about, the intersection of robotics and law and about artificial intelligence and law as well, is the way that it is the way that it is is the way that it strains the law, the way that it challenges the law, the way that it upsets assumptions about the law. Another example is Josh referred to how good this systems are at pattern recognition. They're super good at pattern recognition.

And at one level you think to yourself, okay, well, you can apply that to a bunch of arbitrary domains anywhere that recognizing patterns matters, you could get that. But think about a basic assumption about the law has about privacy. The law assumes that there's things like public stuff like what I'm saying to all of you right now in public or

what you post on Facebook or Instagram or whatever it is that you're using.

These are public things. And then there's the private things that you don't share and you don't want people to know about, right? But what if an artificial intelligence could be designed, and I think it could, they could tell by the very inflection in my voice right now and by the gestures that I'm making and by the way I'm moving my head that, you know, I am a particular thing. A particular thing that I don't want you to know about.

For example, that I know a disturbing amount about my little ponies. Maybe they could tell you that. The point being is that increasingly artificial intelligence is able to derive the intimate from the available.

So does it matter to the law that one day perhaps machines will be like people? No. What matters to the law is, gosh, we can't use this helpful dichotomy between public and private anymore because this technology is eroding our assumptions about it. And so it isn't a sense of very conservative enterprise.

We try to restore the status quo ex ante. We try to think about how legal rights and responsibilities have been disrupted and we try to restore them. But it's a modest project at one level, right? And I think it's going to be, for I think the reasons that Josh thinks too, it's going to be an awfully long time before the law has to confront these deep, deep questions about whether you should treat artificial intelligence as though we're really a person.

So I think I'm going to stop there because I think we really want to have an open conversation. But thank you again for the opportunity. Those are great examples.

Thanks. Thanks very much to both of you. You just reminded me why I didn't go to law school or medical school.

I have the freedom of thinking about these issues without any high stakes attached. So it's an amazing freedom. That said, so I want to get, we already have some questions coming in.

So thank you very much. Here's what I'd like to do. I would like to start because you all invited a philosopher to the party.

So with just some basic kind of conceptual framework and then we'll move into some more interesting questions. But we talk about these artificial intelligence and we talk about emerging technologies. And I'm just wondering if you could say, you've settled a bit.

So feel free to piggyback on that. But what kinds of technologies in fact are you working directly with? So what kinds of systems are we talking about? Both what exists currently and then what do you anticipate coming? If you didn't follow what I said in my talk, it

should be very clear that there isn't actually artificial minds. You guys got that, right? So when we talk about artificial intelligence, I don't think even anyone claims.

I did give a talk in Hong Kong and I found out that there's actually a group there that has a robot body that's animatronic, that's like a chatbot. That they want everyone to treat like a human. It's about the closest you can get.

It's not human. It doesn't have a mind. They somehow want to treat it that way.

I don't understand that. So why should caution you that the question I was going to get at that's come in from the audience is what is a mind and what is consciousness. So when we say there's technologies that lack that, maybe we could.

That's a great question. And I think it's actually maybe more connected to the foundations of law than... Well, we'll talk about that. That's a side topic.

It's a fun one. But really what I'm doing is most of my work is using something called deep learning, which is building neural networks that are pretty complex with interesting structures in a way that allow us to do things that we just didn't know or possible to do with computers even, I'd say even 10 years ago. And maybe even five years ago.

I mean, honestly, it's a really fast-moving field now where I'm constantly surprised about the stuff that's coming out and learning a lot, even though I've been working this field for 20 years now. And it's fun. It's cool.

You find out that the way how it works, it's not actually about this precise detailed coding to get everything precisely right so to work. But it's rather about can you build a structure that has the capability to learn something and then can you get creative about how to apply it in a way that really expands our knowledge. And so that's the type of work I do.

I want to say learn. It is a bit metaphorical. I mean, as exposing it to data, maybe with clever ways of training it to kind of fit all the pieces together and kind of get tuned up so it works.

And then using it to do things like play go or look at a chemical structure to tell you what will happen or to look at a cancer genome to understand which parts of the world are going to be. Or to understand which parts of the mutations are important, why drugs might be useful for treating a patient. Or you can kind of go down the list.

Or facial recognition. We're doing a lot of stuff with looking at pathology slides to understand which kidneys should be transplanted or not. And so those are the sorts of questions.

And we can often get, well I'd say there's a couple of questions here. First, a couple of

things that would come out. One is that we can often get the computer to work as well as a human would and often, very often to work better at a defined focus task.

But it always would be to find focus. If you'd stop making it a defined focus, then that's not the case. The other thing you find out is that it's really hard to get it to deal with these cases where it's not defined.

And the third thing you find out is that especially in high stakes contexts like medical decisions. And cases where understanding is important, like in science, the word science actually means understanding. One of the big issues that we're working through is how to make what the computer is doing or what the machine learning or deep learning is doing understandable.

So that's another question. And I probably said the other thing too is understanding, well okay, so this is clearly a powerful tool. But at times it just can't make the bridge and the gaps.

So how can it work in collaboration with experts? So those are probably like three or four major things that are coming up over over. I want to clarify that the reason that I went to law school was not because I wanted to affect the world. I was not smart enough to go to get a PhD in philosophy.

I want to be clear about that. So I don't work with technology in the same way as Josh does because law professor. But I do have an interdisciplinary lab here on campus that's called the Tech Policy Lab.

And we formally, we have three directors, co-directors. We formally bridge computer science, information science and law. And we have a bunch of robots in our lab and we have a bunch of augmented and virtual reality and so on.

Because I feel that I need to know enough about the technology to not get really, really wrong. You know what I mean? I do spend a lot of time translating. So I work a lot with lawmakers and judges and for a time with the Obama White House, just trying to figure out what should the U.S. policy be towards these things.

And that requires you to understand the technology well enough to be able to talk policymakers who are not familiar with it through it. But that's ends of being the sort of contact that I do. And in terms of my definition of artificial intelligence, I don't disagree that it doesn't have a mind and maybe that's a misnomer.

But I tend to talk about it as a set of techniques aimed at approximating some aspect of human or animal cognition. And so maybe they might be deep learning but it also might be reinforcement learning. It also might be, you know, what they call good old-fashioned AI, symbolic logic.

It's just a way to approximate some aspect of cognition. But again, you can hear again and again, it's a deeply functional definition. And it's like, what are you trying to do? What is this technology afford you? What can you keep ability? And that's my interest in it.

Good. Reinforcement learning is a type of deep learning. In a lot of ways.

Okay, in a lot of ways, but not. Yeah. There are different techniques, right? I mean, one of them and then have been a long around for.

And the other thing, too, Josh, I think it's worth explaining to folks that they don't know. A lot of these techniques were actually developed a long time ago. So we might say that it sounds like AI is moving so fast and we're doing just these amazing things.

We're using techniques that were perhaps developed in the 1950s or in the case of reinforcement learning more likely than 1970s. But what's so different and amazing and new is that we have all this processing power and we also have an unbelievable amount of data. And that those things combine, it's my understanding, those things combine together to make these techniques that were developed a while ago much more powerful.

I know there also have been advances in statistics and other methods that have helped here. But it's interesting to think how you can have these tools that are lying around for a long time. And then suddenly because of different conditions, they become very powerful.

Same. Anyone want to take a stab at what we mean? What if we say it doesn't have a mind? What doesn't it have? I hope you would answer that. I think we think therefore we are.

I think what he's saying there, and you can disagree with a lot of things, but he's describing this reality that you know what your mind is. You have a direct perception of yourself and your mind. You can close your eyes and be thinking about things that aren't here.

You have thoughts, needs, hopes, desires. You have a sense of self. You have a sense of like your past and your future.

You have the ability to think, to feel. It's not that you're emulating that you actually have that. Now when it comes to recognizing that and someone else, it's called the problem of other minds.

And that's a much harder thing. So you have direct perception of your own mind. And then if you look around, and this is actually where it starts to touch on law in an interesting way I would say.

Because you can look around and they have this issue in the philosophy mind called a philosophical zombie, right? There's this possibility that maybe a portion of people out here in this audience. Or all of them. Or all of them.

You are just very good at emulating a mind, but you don't actually have one. Your exact physical duplicates of us, but you have no consciousness. Well, not exactly.

You have to be a doctor. Well, minus one. Whatever consciousness you refuse on.

But you get the point. And so that's a possibility. But here's the thing.

If you don't actually have a mind, that actually changes the moral calculus immensely. Because if someone doesn't actually have a mind, it probably changes our moral responsibilities to them. And so that's actually where it does start to have an impact in philosophy.

He wrote a book that was pretty transformational or pretty important, I think, called "God and Other Minds", right? But the challenge is that you can't actually give, well, we have direct perception of our own minds. And so we can say, "Okay, I have a mind." I can't actually give you a logical reason that's clear, that doesn't have massive loopholes about why you have a mind. And that's a pretty striking thing given that that's a foundation for basic things like in law and morality.

And we don't know why that's the case. I mean, like, how do we all come to the view where we just all believe that what I was buying? People aren't generally going around thinking, you have to go to talk to a philosopher to learn about philosophical zombies, right? And I think what he argued, and I think it's correct, is that it ends up being a proper basically. If it's something that we all believe for no good reason, and it's true.

But we don't actually have a good reason to believe it, and it's true. And that actually works great when we're talking about people. So the problem now becomes, when we start thinking about things, where we don't actually have a proper basic belief about consciousness, and it can start becoming really relevant.

So that was probably a bit more of an explanation than you bargained for, or the person you're question for. But that should help you understand what a mind is, and how it whites so hard, and we don't even have a really good way of accounting for it in other people. Yeah, I like the way David Chalmers describes consciousness is that it's the experiential component of your life, right? It's all the experiences.

So if you think about right now, you've got visual experiences, you've got some maybe tactile experiences, you've got auditory experiences going. It's the inner movie that's constantly playing any time you're awake, or otherwise not unconscious. We might not be able to give a non-circular definition, but that's okay.

So if it's that, then we're wondering whether there is conscious, if experiences can be had in organisms like us, what else could it be had in? Which is an interesting question. Okay, so I think... But the connection to law, though, is especially criminal law. Because in criminal law, we don't just ask for basic ideas of responsibility.

You've got a state of mind. You've got to intend something. And so there's this concept called men's rea, which is that you have the intending mind.

And so in the absence of that, it would be... it really critically matters. And our ability to form a mental model of other people's thinking infuses a lot of areas of the law. So for example, we think about what is reasonable conduct a lot on the basis of our assumptions about people interact.

And so if people... if something's going to happen that we don't know how it works, it can change our calculus. So for example, when hot air balloons were first introduced, it made lawyers really nervous because nobody had expectations about how they would work. You know what I mean? If we had systems that we introduced where we weren't really sure what was going on internally, we wouldn't be able to ascribe viability the way that we do now.

There are so many issues for medicine too, right? Like when we're after deciding, we have a lack of brain functioning of certain kind when you decide, right? That... that noises as such that you don't have the presence of or the potential for regain consciousness. Yeah, I mean, I would also agree with... these are not actually the questions that AI is facing us now. I'd say probably more than these grandfills, south of our questions, which I think are really important to engage.

I think it's... the fundamental questions are in law and ethics right now about what actually does it mean for us to bring this extremely powerful and flexible new technology into the world in a way that's ethical, that we actually know how to manage. That doesn't erode the things that we care about. Good, so let's talk about that.

What kinds of... when you think about the ways that AI is being used for... we'll be used in the... you know, immediate future, fairly immediate future. What kinds of things make you very hopeful about it and what sorts of things give you pause? What do you say? You know, so I mean, I'm aware that I think that the way that this conversation has been marketed and from what I know about who is in the audience, that it behooves us to touch on some questions around faith and what kinds of concerns that we might have. And so one core commitment that I have as a person who has no religious affiliation is that I still nevertheless worry about dehumanizing potential of artificial intelligence and robotics.

And so my little example about the kids being mean to Alexa, you worry about a world in which people are no longer foregrounded. People are no longer the arbiter. And that

folks are getting the social needs that they have met and the cultural needs they have met and other kinds of needs from machines.

And so I worry very much about the potential for dehumanizing effects of technology in the short run. In the long run, I worry about decentering people by imagining that artificial intelligence will let us live forever. Because again, I don't know how many good place fans there are here.

And I don't want to spoil for people who haven't seen it. But a common theme among sort of more secular humanists is that one of the problems of many religions is the centrality of immortality. And the concern is that a belief that you're going to live forever in an unchanging perfect state tends to make your current life and your current time less meaningful.

And I think that could be true both of an immortality that was occasioned by a deity. And it could also be true about an immortality that was occasioned by technology. So in the short run, I think my concerns align very closely with people of faith as I understand from listening what people believe about that, about the dehumanizing effect and the way that it's decentering people.

But in the long run, it's almost like my concerns are also concerns I have about religion. And I'm being a little vulnerable here. I understand that I'm probably among a lot of people who believe very deeply and appropriately so in their faith.

But those are the kinds of concerns I have. I don't know if you want to either of you want to respond to that is. I think what you're saying makes a lot of sense and not regarding me.

I think that I would say I'm a Christian, but even as a Christian, I can see a lot of religion and even a lot of Christian religion can have a very corrupting influence. And so I think that's kind of what you're speaking to. And we're talking about how you're concerned about how we steward the planet.

And like if you're going to have a perfect existence in the future, what do you care about it now? And you're talking about dehumanizing impacts that the possibility that AI could give in a more life. But that means you're concerned about the dehumanizing impact of belief in afterlife from a religious point that you could have. And I think that that's one side of it, which if that's all you have, if that's how you define religion, and if that's a primary way in which you really really religion, yeah, it's going to be dehumanizing.

I just don't think that that's the whole story for what it is and how actually what it really means to be a follower of who Jesus is. And one of the weird things about who Jesus is, is that he actually enters into this world and cares about this world at this moment and cares about it. And we're asked to follow him.

And yeah, so there is an afterlife, but what's actually so interesting about the afterlife that he talks about, and also as you read the whole story of scripture, is that that afterlife is supposed to be here on this earth. It's not in a different place. And so it's not like you screw things up here and you just kind of get a ticket out.

It tells us this idea of like, well, actually, this is the world that he wants to see turned into heaven. And that is actually worth here and coming and seeing so much dignity and value in it that even though it's broken, it's worth engaging. Now, not every Christian approaches things that way, but I didn't say Christians were always compelling.

I said, "Jesus is." And can you at least agree with me that Jesus is greater than Christians? Can I get a name on that? And so, look, I mean, so I get it. So I get what you're saying on that. And I think actually that's actually one of the things that's really, you know, that's really true, but there's also another side of it.

Because I think one of the things that we see when we look at this world is we see a time so much suffering and shortened life that really shouldn't be that way. And that isn't itself just humanizing. And if all we had is this life, that can be to humanizing too.

And so I do think that there's a bit of a paradox. I think that there's something that even if this world gives us a horribly unfair shake, there's still a way that things can be set right. And that is humanizing.

That's also what's allowed people to go willingly in inter-situations that are very self-sacrificing to do very good. That's one component of at least. But I agree to think that that means that that's the only thing that matters is to completely neglect the truth of what we are.

I think that transfer is for this. Does that make sense? Well, if I may, I just want to ask both of you in the audience if you can give feedback about this. Are religious, and again, I'm recognizing both the limits of my understanding and the fact that religious people is hardly some monolithic thing that shares attributes.

It's incredibly heterogeneous. But people who believe in an afterlife of a certain kind, do you find it disturbing this idea that there's a bunch of people who think they want to upload their brain and live forever that way? You see what I mean? Is it going to be like disturbing? Do you find it disturbing? The idea would be there's a door to the right, and the door to the right, which leads to religious afterlife with your soul, and then on the left it's like it's just your brain, and it gets uploaded. It's up to you, which of the two, you know what I mean? I find that a little bit disturbing.

Which part sounds disturbing? Well, first of all, it feels so faithless. It feels so denying a faith to say that people have to take this whole immortality thing into our own hands. Do you really mean it's just sort of like, you know, we've been believing this thing about

being immortal because like, you know, God said... Oh, you're saying that the same thing? Yeah, so I would say that.

In the conflict they are right. I mean, I'm genuinely very, very curious. Well, there's been a lot of Christians writing about this thing called transhumanism and the singularity.

Yeah, yeah. And that's exactly the negative view they take of it. Okay, okay.

I take a little bit of a different view. So I'm probably a bit more of a... I'm a friendly guy, right? I like to see the difference. So I kind of see that and say, wow, that's just so clearly a crazy thing to want to do.

Like, just think about it. Just think about it. Like, let's say we can download our brain into a digital thing.

And so these people are even thinking about like doing destructive stuff where they like slice their brains and tiny pieces and put it in. Now, that might be a good copy of my brain, but is that making that copy of my brain worth actually destroying my brain? That sounds crazy. Okay, so that's the first thing.

I have to use that technique. But if they use the technique that left your brain just fine, it made like a weird copy of it that could persist forever. Yeah, but now we're in.

Yeah, I mean, actually this whole thing is in science research. I said, now, okay, we're already in science research. But then... I think what's... Yeah, so it is like this... it is like a religious belief.

Because it's not random in reality. It's not even clear why this makes sense. But it's just so... what I find so interesting about it is it's not actually really logical, but it's speaking to something like, what is it about us as humans that make us think about that? And I think that's really valuable to understand.

I think we'd like to have continued experiences, right? We have kind of an interest in the continuation of our experiences, provided they're like marginally good. Yeah, that's right. Because I don't want to say any other experience.

You don't want to just any other experience. Well, I think it's interesting about it. Okay, so here's the thing.

I think they're mainly atheists that are wondering this way. Mainly. Not always.

But they're all not... or nuns. Nuns is the right way to put it. Or agnoviligists.

But here's the thing, like that... I thought you said nuns too. I don't think nuns want to... sorry. Nuns means they're not any individual religion.

Not any individual religion. Oh, you thought I meant N-U-N. That didn't make a lot of sense.

And O-N-E's. Alright, cool. Alright, so they're the nuns.

Non-fist. Yeah, now I lost my point. But what's interesting about it is that... that there was one thing that seems like everyone should really understand is the inevitability of death.

That somehow we have this impulse to want more than that. It's weird to think that that would happen. Like, why is it that we have this impulse, even if we're not religious, to desire something that may not even be good? It doesn't be sensible.

I think that that's a really interesting puzzle. And like the historical Christian answer is that God put eternity on our hearts. That he actually made us for that purpose.

And so we have a craving for something. We have a longing for something that we've never experienced yet. And actually don't even have a good account for that.

It could be impossible. And that's actually a clue. It's like a signpost that there actually might be something there.

And that's even like an account for why there are other religions and things like that. I would say like we could tell a different story where it's like a survival instinct and then the religious folks tell the story as they say. Oh sure.

None of this is proof. I'm just saying. Alternative explanations here.

The other thing I think about is like... and this I think is a bit of a caricature and a cliche as so many of these things are. But I think about there have been cultures and there are cultures that seem to imbue some or all objects with a sense of a spirit and a sense of... You know what I mean? And I wonder how that interacts with our existing tendency to anthropomorphize. Right? And so it's not necessarily dehumanizing but it's destabilizing.

It's destabilizing to that way of thinking to have objects that you have as a category. You think this object is a thing but it has in a sense its own... Not a soul precisely but I'm thinking about animism here. You know what I mean? It's imbued with some kind of spiritual content.

And then you have this other thing over here which is imbued with the spiritual content but also emulates us. You know and looks like us and maybe maybe one day we'll make claims to be like us. Right? And so what is the content of their spirit? Anyway, the point is that it much like it has a destabilizing effect on law, it feels like some of the more outlandish forward looking theories about what AI can do has a destabilizing effect on spirituality.

Well yeah, it brings us to this question about what it means to be human. There was a statement, an Evangelical statement on AI that was put out early last year. And I wrote a response to it that I was really fortunate to get published in the Wall Street Journal.

And it was interesting because the reason why I was actually disputing the mode is that they kind of put forward like what they thought Evangelical should really be saying about that AI. And some of the stuff is actually good and I would agree with it but there are certain points like, "Oh wait a minute, I don't know about that." One of the points they said is like it's kind of like a statement and a starting point was that artificial intelligence could never be in the image of God. And so this is an interesting question actually.

And I had the privilege of sitting down talking to a bunch of theologians right when this thing came out, seeing how talking came about it. These are conservative theologians too. And I said, "Well okay obviously there is no actual strong AI that actually has a mind, what do you think? Do you actually think that AI could ever be in the image of God?" And that's an interesting question because right now you look around the only things that are in the image of God are other humans, meaning other homo sapiens, other biological beings.

And it was really interesting to get to the responses. They all said, "Well actually scripture doesn't tell us. Maybe God could give it the image of God in the future." Like who is to say that God couldn't actually put a soul within a machine.

He could do that. And so they were saying, "Well actually I guess we don't know." And that does destabilize theology, but in a very interesting way to stabilize it in a way that creates space for conversation. If you actually look at what the image of God is in theology, I think the best high level explanation that's brief, which is actually true, is that it's just been this historical contemplation for over 2,000 years now, or a long time, about how do we think about it? How do we explain? How do we discuss what it means to be human? Is it located in our relationships, our attributes, and our actions? Is it, what's the essential part of what it means to be human and what's the contingent part that might have been added by the fall? How do we actually live as truly fully human people? Is the question that arises when we think about the image of God? And it has had a huge impact on society too when you think about people like Martin Luther King, right, who talked about human rights and dignity connected to the image of God as well.

And all those things, when we think about the possibility of other minds, whether it be artificial intelligence, intelligent aliens, or people who live before Adam and Eve, these are all the sorts of questions that this brings us to. And all those questions are those people, all those other minds bring us back to this question about what it means to be human. Yeah, I think we've used the word dehumanizing, and I think it's really interesting because we say that we clearly have some idea of what it is to be human, or humanizing, right, where we would be the opposite of dehumanizing.

But I think there are sometimes we care about what is to be a human person, and sometimes we're just interested in what is to be a person, so what kinds of things could count, especially as we start talking about what sorts of things could be right there. See, the law thinks that Microsoft is a person. Well, they've already gotten so far that doesn't have anything to do with that, right? Like a corporation is a person.

Well, that's what misunderstands. Oh, I understand. Yeah, they understand.

Okay. It's really so again. I know you're like, that is not a person, is it? Yeah.

So we talk about this. Again, it's kind of a caricature of the law that the law is just like, "Who ever is a person?" I don't know. You're a person? Come on.

It's like the Oprah Winfrey Show, where it's like, "You're a person, and you get a person, and you get a person." That's not how the law actually works. What happens is that the law recognizes that because corporations are collections of people, it will afford that corporation's certain rights and liability selectively, right? That are parasitic upon the fact that corporations are created by people. And there are all kinds of complicated reasons, but corporations only get some sets of rights and not all rights.

And the rationale is pragmatic and consequentialist, not because intrinsically they're worth it. And the same thing has been happening with artificial intelligence. So there's a literature around bots.

I don't even know if bots are artificially intelligent. But these bots, like these are these software agents that can go on Twitter and whatever. They interact with you.

They're trying to pass the Turing test or whatever. And there's been some really sophisticated conversation by law professors about whether or not bots should be entitled to free speech rights. And they have argued that largely they should.

And the reason that bots should get free speech rights, the argument goes, is because we as humans have a right to listen to them. And we should not allow the government to selectively tell us which messages we should hear or not hear. And we should not cut off a mode of communication that, of course, you can be traced back to humans.

It's not that bots become people suddenly, and yet, bots have this right, this right to speak that is maybe parasitic upon our right to listen or parasitic upon the idea that they're speaking at the instigation of people. And so, again, it's a very pragmatic sort of discourse around the law. It does retreat to first principles sometimes, and certainly it cares about ethics.

But it's very pragmatic. Right. Right.

Okay. I'm going to get to some questions that you all asked. Okay.

And there are some specific ones here. So Professor Kalo, do you have an-- oh, what do you think? Oh, what do you go? There it is. It flipped around.

Sorry. Do you have an opinion about who is liable for the decision-- sorry-- of an intelligent computer that causes harm. So in the case of self-driving.

Oh, and not only do I have an opinion about this, I have three or four articles about this topic that I won't-- but I won't get into it. So, you know, I'll tell you, it's fascinating to think about the conditions under which there would be liability for a robot or an artificial intelligence. Usually the law would not struggle very much because it would say, hey, you built this driverless car, Uber, and your driverless car ran into somebody, and so you're liable Uber, right? And in fact, when the driverless car in Arizona struck a pedestrian who was crossing the street, it wouldn't have taken a court very long to figure out that Uber was responsible for that.

And that's because when you make something and it foreseeably causes a harm, you, the creator, are liable for it. Where it gets really, really tricky is where robots do things that humans did not expect. That is hard, okay? That implicates something that's called proximate causation, which is the idea that we hold people responsible for the foreseeable negative impacts.

And we think of foreseeable as foreseeable category of harm. So I'll tell you really quickly a quick example of this. I hope I can get this out fast.

But imagine that a group of engineers build a driverless hybrid car, and in addition to driving itself, the car is instructed to experiment with fuel efficiency, and so it can't violate the law, but it can experiment with battery versus gas or whatever in order to optimize for efficiency using various techniques. And they build this thing, and they're really happy with it, and it goes out. And this car over time comes to learn that it performs better from an energy perspective.

It's more efficient if it starts the day off with a full battery, okay? It just figures that out. Like whenever I start with a full battery, I have a better day. Great.

Tonight, the family that owns the car forget to plug it into the plug, and it's in the garage, and the car's like, "Well, I'm not plugged in, and I'm not going to have a full battery. No problem. I'll just run the gas engine." Killing everybody in the house by asphyxiation.

You go to those engineers and you say, "You're liable. You killed some people with your machine." And they'll say, "Not only did we not intend for this to happen, but we didn't have any idea that this would happen. We didn't predict this at all.

We didn't even predict that asphyxiation was a way you could die with this car. If we'd run over somebody or driven off a cliff or something like that, I get it. But we had no idea

that this could happen.

That's a genuine, difficult car. But now they have constructive notice. Now they do.

But does that's no comfort to the survivors of the family? So the truth is that another example is when Microsoft built-- What I mean is your talk is giving them constructive notice for the future. Anybody in this room makes a hybrid, intelligent car. You are not off the hook, because you heard Ryan say.

They should foresee it now, man. You should foresee it. Because of your imagination.

Indeed. Indeed. But whenever machines-- so machines are interesting to the law in part because they display emergent behavior.

That is to say they do things that-- what's great about machines a lot of the time is that they can solve problems in ways people wouldn't. If people would solve the problem that way, then they should just do it. But the machine-- you know what I mean? The machine doesn't.

That's why so many people have learned so much from AlphaGo because the machine plays Go in a way that the very best players did not envision same with AlphaZero and chess. It's really changed the way people play chess because it's solving a problem in a new way. So it displays this unpredictable behavior and that's its benefit.

But that's also its danger. And so for example, when Microsoft built TAY, which was this bot on Twitter, and it very quickly-- and it's supposed to like emergently engage in dialogue. And it very quickly turned into a racist troll and started to do horrible things.

I mean, just horrible things. And ultimately, Microsoft had to turn it off. And then, by the way, a little while later, they turned it back on for like a second to be like, you still racist? I'm still racist.

I have to turn it back off again. Anyway, but Microsoft TAY said a bunch of things that you cannot say in Germany. You know what I mean? Like hate speech law is such that the kind of stuff that TAY was saying is illegal in Germany.

And does that mean that Microsoft should be hauled into a German court and forced to pay for denying the Holocaust to happen? You know what I mean? It's just like-- but no, because the law will look and try to determine whether Microsoft intended this or at least were sought to behavior and so on. So it provides a pretty difficult challenge. And so that's a great question.

I appreciate the opportunity to answer that. Okay, here's another one. What happens if an AI becomes religious? If they have human traits and abilities-- and I take it functions-- then theoretically, they can adopt faith as well.

Would that be legitimate in a church? Have you guys seen Battlestar Galactica? I feel the pews. Yes! That's like a great show, right? And what's going on? There's silons, right, that are artificial life, and then there's humans. And incidentally, the story actually ends up following the plot of my book, The Genealogical Adam and Eve, which has to take a look at it.

It's kind of funny how it works out. That was accidental. But anyways, getting back to the story.

So the weird thing is that all the humans that we think about them as humans are all-- they're pagan. They-- or not pagan, they are like all fall like the Greek gods, a whole bunch of gods, a polytheist, right? But then the silons are all monotheists. They're talking about the one true god.

It's interesting, right? It's the religious. So they don't think that humans, even though humans created them, are their gods. They think that there's another god.

Yeah, so I don't know. I mean, I think it's one of those questions that if we ever saw a-- I'm not talking about a robot body that's animatronic, like the Chuck E. Cheese thing. But I mean, if we actually saw something that we were really puzzling if it was a mind, and one of the emergent behaviors seemed to show was it was actually showing up in church and doing things, that'd be really interesting, wouldn't it? I don't know what to make of it.

I was talking to a Catholic theologian recently. So I wrote a book-- the book I wrote is actually about human origins. And the big idea that I was really pressing Catholic theologians on is what about rational souls outside the garden? Is there anything really that rules us out? And I think it's an open question in Catholicism.

And if that's the case, it just really creates some interesting questions about people, or ex-amni, the distant past. Not today, everyone today descends from Adam and Eve. So it doesn't go down the polytheist route, or polygenesis route.

But regardless, he told me-- well, he completely didn't like this idea, this particular-- a Catholic theologian. And he said, you know, but let me tell you about the slime man. I saw a man like a rise up out of the slime, and then kind of came up, and I could actually talk to him, and he asked me to appetize.

I think he should be baptized. I think we should give him a communion and do all that. Which is really strange, because that slime man, as he was describing, wouldn't be a descendant of Adam and Eve.

And so what he was saying is that as a theologian is just that ready, direct appearance of a mind was enough to grant that personhood. And full ability to access the Catholic sacraments, even, even though he completely denied the idea of rational souls outside

the garden. So it's just kind of something that would be just so destabilizing.

But if that happened, that's what is responsible, which I thought was interesting. Does that make sense? Well, you mentioned before the idea that maybe we'll encounter like a machine mind before an alien mind. I mean, if we did have an alien species come to visit us, I don't see why they couldn't convert to whatever religion they chose.

Right? And so anyway, yeah. Yeah, so there's a great article by C.S. Lewis called Religion and Rocketry, where he kind of wonders about the possibility of intelligent aliens and what that would mean. Because like, as a Christian, we would believe that whatever aliens are out there, they were created by God too.

But that raises questions. Well, they can't be fallen in Adam's sin, can they? And that means that Jesus is our God too, but in what way? I mean, there's just a lot of really interesting questions. I mean, did Jesus go and incarnate over there too? Well, this might be too part of why personhood might matter more than humanity or being human.

Because I mean, I understand an octopus is very intelligent, highly intelligent, but does not anything like a brain structure of the sort that we have? So I mean-- But is it a person? Oh, I would think that is very much a candidate for being a person. A candidate, but that's not an answer. So there's what is, and there's can I tell? Right, or what do I know about it? Well, do you think that octopi are people? Like people.

Do you think that personhood? Is it just a plural person? Yeah. Yeah. Do you think that personhood? Sure.

Sure. Yeah, I'll just go on record. Yes.

I don't know. I don't know. I think, yeah.

Dolphins. I think I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. I'm going to go on record. I'm going to go on record.

I'm going to go on record. Well, it's an interesting story because you're kind of identifying the problem because you don't want them abusing the similitude of a human. You also don't want them to treat the similitude of a human as human.

That's the fundamental thing. What you're saying is you're thinking about this with your wife and with your kids and you're wondering about the dehumanizing impact such as the point that you're even going to take that out of your life right now because you think it's better not to have that thing. What I want to actually suggest is that maybe Alexa was a humanizing impact because it actually brought you to that grand question as a family.

I might think that you should have kept it around, but actually that had brought you to the question and that you actually handled it in a deeply human way. That may be actually what the right thing is to emulate. What he did there of realizing when all this stuff is happening that we are supposed to be engaging those deeply meaningful questions about what does it mean to be fully human? How do I actually do that? How do I not be dehumanized by a deeply dehumanizing world? How do I not dehumanize others? How do we really become that? In fact, I think part of that is even having challenges to our humanity that we think deeply about and engage in that way.

I thought that was a great example. Maybe there's a flip side to it. Maybe Alexa actually was a deeply humanizing fear of kids to see you think about it.

That's interesting. It certainly was a provocation for sure. I would just say that, look, I have no moral authority.

I'm just an individual speculating about the world. What I would say is that these are disorienting times. Make sure that you keep your values close and that you don't let these disruptions, make sure it's important to question your beliefs and it's important to question your values.

Ultimately, we have a lot of agency. I think one of the real problems with technology, especially the internet and now artificial intelligence, is that it's something that's cooked up by other people and then kind of voiced it on us in a way that we don't have a lot of say over. Then after it's been distributed, suddenly it starts to pose these quandaries and erode these values and channel our thinking and our beliefs and so on.

The great thing about being people is that we get to decide the kind of world that we live in. I encourage you to push back hard when you feel like technology is encroaching on things that you hold dear because now is the time to do that and it will be too late if you don't do it now. If you like this and you want to hear more, like, share, review and subscribe to this podcast.

And from all of us here at the Veritas Forum, thank you.

[MUSIC]