

## Is the Human Brain Just GPT-3 Made of Meat?

<https://mindmatters.ai/podcast/ep223>

Announcer:

On this edition of Mind Matters News, we listen to an interview with Robert J. Marks on the podcast Philosophy for the People. The host is Pat Flynn. Let's listen in.

Pat Flynn:

Okay, everybody, welcome back to the podcast. I am joined today by Dr. Robert J. Marks, and we're going to talk about whether or not you are a computer. AI has kind of exploded recently. I mean, it's been talked about for a long time, but it seems like there's just been this incredible leap forward, especially when it comes to chatbots. And, Dr. Marks, I had some recent AI images of myself generated. I should have brought them on for the conversation because some of them were pretty good and some of them were crazy off the mark.

Robert J Marks:

So, how did you get AI images of yourself? What software did you use?

Pat Flynn:

There is a new app, if anybody is interested, and I can upload one as we're talking here. It's called Lensa. And what you do is you upload about 10 to 20 selfies of yourself. In fact, if you see the background, one of them is there. If you see right above my head, that is an image of me that was generated by this app.

Robert J Marks:

Wait a second. I see in the upper right I think a picture of Aristotle or Plato.

Pat Flynn:

Go to the other side. Go to the other side. It's more in the background.

Robert J Marks:

Okay. The other side.

Pat Flynn:

See the glasses up there?

Robert J Marks:

Oh yes, I do. I do. Okay.

Pat Flynn:

That is from that app. You upload 10 to 20 selfies and then the AI works its AI thing on those images and spits out however many images you're willing to pay for. It's a paid app, so it's not free, 100 or 200. And some of them are, I have to say, some of them were pretty impressive. Some of them were like, "Okay, I

don't know how this came about." Anyways, point being, it seems like with the chatbots and AIR and AI music, people have been paying a lot more attention to AI. So I thought, "Now is a great time to have a conversation about this," and Dr. Marks is the resident expert. He's got a book out, a great book called, Non-Computable You, and it's on-

Robert J Marks:

And you read it, right? Or you read part of it?

Pat Flynn:

Yes, I did. Yeah. And I forgot how I discovered it, but I discovered it independently of our mutual associates and then I realized we had mutual associates. I'm like, "Hook me up with Dr. Marks. I want to talk to him. This is a good book."

Robert J Marks:

Okay, that's good.

Pat Flynn:

So, I'm sorry, I'm getting ahead of myself. Dr. Marks, if you wouldn't mind, give us a little bit of background about yourself. Who are you, what do you do, and what got you interested in all these kind of wider, really philosophical debates about the human person and AI and all that?

Robert J Marks:

Well, I've been involved in artificial intelligence for over 30 years. So I've been publishing and I still publish in the area. I go out in archival peer-review journals, and I finally decided that there's so much going on right now with the hype associated with artificial intelligence. The idea that artificial intelligence is going to take over like it did in the Terminator movie, or it's going to take over like in the Matrix, and we're all going to be lying in bathtubs full of goo with wires connected to us. And I wanted to do pushbacks on that. It turns out that artificial intelligence is incredible, and I don't want to diminish the things that artificial intelligence can do because they can do remarkable things.

However, there are certain brick walls that artificial intelligence will never go through in terms of duplicating what humans do. And there's a lot of proponents of this. We have people like Ray Kurzweil in this book, The Singularity. He wants to come out and do... Well, he believes that, okay, that humans have evolved to the point where we can't evolve anymore, so we have to switch modalities. So in order for humans to evolve to the next stage, we have to go to silicon, and we have to program ourselves in silicon. And then we are going to have AI that writes better AI, that writes better AI, reaching something which is called artificial general intelligence. That's going to keep on and eventually become a super intelligence, which is going to dwarf us all intellectually and make us all slaves and pets and things of that sort.

Pat Flynn:

And we've seen the Terminator, right?

Robert J Marks:

Yeah, exactly. So I wanted to push back against that because this is science fiction and it is so far from some of the things that are presented in both philosophy and computer science that show again that

there are certain aspects of artificial intelligence that will never be accomplished. Never be accomplished by artificial intelligence. So humans are fearfully and wonderfully made, and we celebrate the idea of exceptionalism in these certain areas. Now, there are certain things computers do that they do a lot better than we do. I think they can add numbers and multiply numbers a lot better.

Pat Flynn:

Better than I can. That's for sure.

Robert J Marks:

Better than I can. So, yeah, it's not saying that everything can't be done better by a computer, but as far as duplicating certain aspects of human beings, I would say, for example, the simple emotions, love, empathy, hope, faith. These are things which will never be duplicated by humans. But if you drill down deeper, there are aspects like creativity, there is understanding. Computers will never understand what they do. They can add the numbers three and six, but they don't understand what the numbers three and six are. And they will never be sentient. You have to define these terms, by the way, before you can talk about them, because people use seductive semantics and they go in with these fuzzily defined terms. Well, if you have a term which is defined in a fuzzy way, you can make any case that you want. So I'm...

Pat Flynn:

Like intelligence and like creativity, right?

Robert J Marks:

Exactly. Exactly. Yeah. I think Kurzweil says, "In certain year computers are going to be a million times..." I forget exactly what the figure is, but they're going to be a million times more intelligent than human beings. Well, how the heck do you measure that? I mean, what does that mean? And in terms of multiplication, in terms of things computer programs currently do, yeah, they are millions of times probably more intelligent than I am, but they're never going to duplicate some of those human aspects that I mentioned.

Pat Flynn:

Yeah. Where do you think is one of the best places to start with framing this conversation? An analogy I heard one time I thought was always very interesting. It came from a philosopher named David Bentley Hart, and he argues against strong AI, and I think he'd be on the same page as you are, Dr. Marks. And one of the things he said quite poetically is, "When we first started building computers, we kind of used the analogy of ourselves and put it onto computers and saying, 'Well, it's thinking. It's doing all this.' But then we forgot that that analogy came from us and now we're taking it back from computers and reapplying it to us thinking, 'We're just computers.'"

Robert J Marks:

Exactly.

Pat Flynn:

We say, kind of like a historical confusion or misremembering that went on. And I figured, that actually sums up a lot of the issues in a funny and interesting way, and I'd be curious if you agree with that.

Robert J Marks:

Well, let's see. So, boil down your question again. That was a one question.

Pat Flynn:

Yes. So, I think one of the things that I've always been interested in and would want to emphasize is that the computer, in and of itself, is running on logical operations, but the computer, in and of itself, is not in and of itself carrying out logical operations or processing informations, any more than a piece of scratch paper is, right? Because considered by itself-

Robert J Marks:

Oh, exactly.

Pat Flynn:

... apart from our conventions, our intentions, it's totally meaningless, right?

Robert J Marks:

Yes, yes.

Pat Flynn:

It's totally meaningless. So there's always this background that if we forget about the background, we're going to read more into it than is actually there. But it's all observer relative in a sense, isn't it? Right? You see what I'm getting at now?

Robert J Marks:

Yeah. There's a number of things you mentioned. I have a bunch of different thoughts on that. One of them is the head of the Allen Institute for Artificial Intelligence in Seattle, and he said... His name is Oren Etzioni, I believe is how you pronounce it. And he said, "Artificial intelligence is nothing more than a pencil," in the sense that if you look at a computer program, you are telling that computer program what to do step-by-step-by-step. And if you had a time, it might take you a billion, two billion years to do it, you could go ahead and do all of that with paper and pencil. So, yeah, you're right, you're right in that sense.

But as far as whether we are... Well, let me interject. There's a great story by Emo Philips, my favorite comedian of all time. He said, "I was walking one day and I realized that my brain was the most important, remarkable organ in my body. And then I realized, who was telling me this? It was my brain that was telling me this." So there's a little bit of deception here. So, let me change the topic a little bit. We do have something called the mind-body problem, which I'm sure that you're familiar with, comes in the area of the philosophy of mind. And that is whether the mind is something which is an outgrowth of the brain, or whether the brain and the mind are some way dissociated with each other, and whether the mind is external in some sense from the brain.

And this has been going on for years and years and years, and I'm sure you're familiar with some of the old debates of Descartes. I think he was one of the early people that talked about the mind-body problem, and he referred to the mind as the soul. And that was interesting. But the question is, which one you philosophically agree with? What's your worldview? What's your ideology? And if you have an ideology that we are computers made out of meat, then the only thing you can do is say, "AI must be

able to duplicate it." After all, our brains are just computers made out of meat, so therefore we can go ahead and we can simulate and duplicate the human mind with a computer.

But the dualist would argue that the mind is external from the brain. There are certain things that the mind does that can't be explained by the brain. I was influenced a lot on this by Roger Penrose. He won the Nobel Prize a couple of years ago. He wrote a great book called *The Emperor's New Mind*, where he convinced me that some of these human attributes such as creativity were beyond the ability of humans to do. The basic idea-

Pat Flynn:

Beyond the ability of computers?

Robert J Marks:

Yes, I'm sorry. Beyond the ability of computers. Thank you. Yeah, I misspoke there. And so he makes the point that back in the 1930s, Alan Turing, who was one of the founders of computer science, showed that there were things that were non-algorithmic. An algorithmic is a step-by-step procedure to do something, and all computer programs, including artificial intelligence or algorithms, if you have artificial intelligence, is an algorithm. It's a step-by-step procedure to do something. So the question in computer science was, are there things which are non-algorithmic? Are there problems which a computer cannot solve? And Turing was the first to show the first problem that was not solvable by computers.

It was provably non-computable. This wasn't conjecture. This wasn't philosophical. It was a mathematical truth. Since that time, people such as Gregory Chaitin and Rice have generalized and shown that there's a number of different problems that can't be solved by a computer. They are non-algorithmic. So this begs the question, and this is the premise of my book, are there aspects of humans which are non-computable? Because if they are non-computable, that means they're beyond the reach of a computer, they're beyond the reach of an algorithm. Yes. And I think we're starting to accumulate evidence towards that. Stephen Hawking said in his book. I love this quote. Stephen Hawking was an atheist, and he would've won the Nobel Prize with Roger Penrose. They worked together on black hole sort of stuff.

But Stephen Hawking said that, "Nothing in physics is ever proven. All's you do is accumulate evidence." And so that's what we're starting to do with the mind-body problem, is we're beginning to accumulate evidence that indeed there is, or there are, attributes of the human which are non-computable. And the non-computable aspect is very important because it means if it's non-computable, it can't be computed today, can't be computer tomorrow.

Pat Flynn:

In principle.

Robert J Marks:

A lot of people say, "Hey, hey. No, it can't be." Not with algorithmic computers. Not in principle. It absolutely can't. Because it's non-computable, because computers of tomorrow will still use algorithms. They will still use something which is a step-by-step procedure. And if it's non-algorithmic today, it's non-algorithmic in the future.

Pat Flynn:

Yeah, and that's a really important point. And I want to explore all those issues including mind and brain. Soul and brain. I actually don't think the mind is a soul. I think that's a mistake, but that's another issue entirely here.

Robert J Marks:

I'm sure meanings of words have changed over the years. That's what they currently referred to it as, but you're right.

Pat Flynn:

Yep. So, one is like drawing lines in the sand, because there were a number of people who said certain things wouldn't happen. They did happen and they had egg on their face. Think like chess, right? It's like, "Oh, no, a computer will never beat a human in chess or his games..." Ah, no, it kind of whoops up on us now, doesn't it, right? Or it'll never be able to generate art or poetry or stuff like that. And I don't know, these chatbots are getting pretty impressive.

So figuring out where these boundaries are and being able to give a case for them being principle batteries such that if they were crossed, you would have to put your hands up and say, "All right, I was definitely wrong about that." I think it's an important thing to do and showing why those boundaries are different than, say, the success in chess matches or getting Pat Flynn's ugly mug in some sort of special outer space rendering or something like that. You see what I'm getting at here, Dr. Marks-

Robert J Marks:

I do. Let's boil that down. You actually raised three topics here. Let me address them one at a time.

Pat Flynn:

Sure.

Robert J Marks:

Number one is chess. Chess is algorithmic. Claude Shannon, back in the '40s, came up with a technique to do the next best move in chess. You wanted to make the move which minimized the maximum damage your opponent could do. You wanted to make a move that minimized the maximum damage that you could do. And this again was back, I believe, in the '50s, and he actually had a very small chess board. He wasn't able to do 64 x 64. But using reinforcement learning, we know now that that can be done.

So the question you have to ask about these things, are they algorithmic? Is there a set of rules which capture and define the process that you're talking about? In chess, yes, they are algorithmic. So I don't think it should be surprising that artificial intelligence has overtaken chess. I also don't think it's surprising that AlphaGo, when it beat Lee Sedol, the world champion in Go, that was algorithmic also. And so, no, I guess I'm not surprised at those things. I don't think that computers, for example, have ever been creative, if I could define creativity.

Pat Flynn:

Yes, important-

Robert J Marks:

Creativity is best defined by Selmer Bringsjord at Rensselaer. He used something called the Lovelace Test. Ada Lovelace was the first computer programmer of the 19th century. She worked on an invention which was conceived of, but never built by a scientist, a genius named Babbage. Anyway, the Lovelace Test asked the question of artificial intelligence. Does the artificial intelligence do something which was beyond the intent or explanation of the programmer? Does it do something which is beyond the intent or the explanation of the computer programmer? And that is going to define creativity. And to date, I know of no artificial intelligence that has passed the Lovelace Test for being creative.

This is incredibly important because one of the things that it debunks is this idea of computer programs writing better computer programs, writing better computer programs, because a computer program writing a better computer program preassumes creativity. That better computer program has to be beyond the intent of the original programmer in order to be creative. And there's been no demonstration that that has indeed happened. And in fact, I don't believe it'll ever happen. I don't think the Lovelace Test will ever be accomplished. You talk about computing art and things of that sort. If we could talk about that for a second.

Pat Flynn:

Yeah, please. Yeah .he music that it's turning out now.

Robert J Marks:

Oh, exactly. But let's talk about a scenario where you want to write artificial intelligence that composes Baroque music. What do you? You gather a bunch of pieces written by Bach and Handel, people of the Baroque period, and you feed it in some way to the artificial intelligence. And so the artificial intelligence looks at it and scratches its algorithmic head, and then it begins to generate music. What does that music sound like? It's Baroque, isn't it? It will never generate music that sounds like Wagner or Stravinsky or jazz. It will not be creative at all. It is constrained to live inside of the silo of its training data. It lives inside the silo of its training data.

There's the old saying, "You have to think outside the box." Well, the training data of artificial intelligence defines the box, and in order to be creative, you have to think outside the box. And the interesting thing is, the creativity normally requires you to discard things inside the box. Discard the dogma. If you look back at some of the great advances in science, for example, they had to discard the dogma at the time to go forward. A classic example is Einstein. Okay? Used to figure that the speed of light was relative to the velocity of the observer, the relative velocity. No, the speed of light was absolute.

They also used to believe that there was something in space called ether because they knew sound waves had to travel through solids. And so, therefore, electromagnetic stuff had to have something that it traveled through, and that must be, I don't know, let's call it ether. And the Michelson-Morley experiment kind of debunked that. So if you asked the vast majority of scientists during Einstein's day, inside the box, about the relative velocity of light, about the existence of ether, they would've agreed with that dogma. But in order to be creative, Einstein had to step outside the box, outside the box, discard the dogma, and do something which turned out to be creative. So I maintain that, yeah, artificial intelligence can generate jazz, but guess what it was trained with? Jazz. Right?

Pat Flynn:

Right.

Robert J Marks:

Your picture on the upper left here, I think it's on-

Pat Flynn:

Here. I actually got one right here. There's another one of me.

Robert J Marks:

Oh, my gosh.

Pat Flynn:

From AI, right? Not terrible. I mean the right eye is a little... I don't think I'm that squinty, Dr. Marks, but...

Robert J Marks:

No. I tell you, you look like a superhero there. Or-

Pat Flynn:

Yeah, I think that was the package I paid for. I wanted it to make me look like a superhero.

Robert J Marks:

Is that right? That's a great Christmas present. What was the name of that software again?

Pat Flynn:

It's called Lensa. I recommend it, especially for podcasts like this, it helps you come up with little avatars and catchy stuff.

Look, I'm a fan of all the cool stuff that it can do. L-E-N-S-A. L-E-N-S-A.

Robert J Marks:

Lensa. Oh, that's a great app. Okay. You want me to talk about chatbots a little bit?

Pat Flynn:

Yeah, because-

Robert J Marks:

Yeah. Okay, chatbots-

Pat Flynn:

... people are really being impressed by these new GPT...

Robert J Marks:

It's incredible. And I actually purchased GPT-3 and I sat down and played with it for a while. It is truly remarkable. There's a few things I learned right away. Number one, it doesn't keep up with history. I asked it who the president of the United States was, and it said Donald Trump. So it's a little bit behind the times, and it gets a lot of facts wrong. It isn't a place to go for facts. But here's the way the GPT-3

works. You have something called syntax. Syntax is the relationship between words. If you have a word, what words kind of belong with it, and they looked at this juxtaposition of the different words, and they learn how these words flow. And they train on billions of different articles, all in the syntax. But nowhere, nowhere do they talk about the semantics of what's going on.

They don't talk about the meaning of what is going on. And what's interesting about this, since GPT-3 has been trained with all of Wikipedia, there's a number of articles in Wikipedia that explain things like simple arithmetic, a long division, and even how to take square roots. There's probably articles. GPT-3, in reading those, had no idea of the instructions. It does not learn from the instructions. All it does is look at the relationship between the words. And because of that, the idea of simple arithmetic has escaped GPT-3. I asked it, for example, what was two times one, one, one, one, and the answer was 22. And it does get some simple arithmetic right, because I suspect that somewhere in the training it got, it's seen that simple arithmetic, but it's never learned elementary arithmetic.

Now, the interesting thing is I asked Alexa these same simple math problems, and you know what? Alexa got them right. And do you know why it got them right? Because Alexa was not trained on the syntax. I think, just like Google, when you ask it a math problem, it switches modes. It goes over to a math mode, uses something totally different and answers the question. So can GPT-3 be trained to do mathematics? Yeah, absolutely. But currently it doesn't do that. Another thing that syntax doesn't do is learn about self-reference. In other words, a sentence referencing itself in some fashion. Like I asked it, "Does this sentence contain 10 words?" Okay. So that was something that a six year old can answer.

"Does this question contain six words?" Now a child would've got that, I think. Anybody that knows how to count. But this was difficult, maybe impossible, for GPT-3 to learn from syntax alone. It did not learn self-referencing. It did not learn elementary arithmetic or elementary algebra, even though it had all of these tutorials instructing it how to do that. It had no idea what the meaning of those tutorials was. It was totally looking at the syntax. But you're right, it is incredible and it does wonderful things. I don't want to diminish GPT-3. It is an astonishing piece of software and can do remarkable things, but it is limited. And so the bottom line is, does this pass the Lovelace Test for creativity? No, it doesn't. It was trained on syntax and it does stuff which is good with syntax.

Pat Flynn:

So I want to jam on a couple of themes there, because you brought up the notion of semantics or meaning. And one of the things that really shifted my thought in philosophy of mind is that, Dr. Marks, when it comes to us and our thought, certainly our formal thinking, we have determinate semantic content. We really know what we're thinking about. We understand things like triangularity and modus ponens and so on, like that. But if you just look at-

Robert J Marks:

All in the world of Plato, right?

Pat Flynn:

Yeah, in a sense. I mean, I'm making an argument here following Aristotle and many other people, that it's there, that the intellect is actually immaterial because any material thing is particular, and it's always ambiguous into how it could be interpreted. It can approximate, but it can never be adding versus shm-adding, for example. How can know, determinately, what's going on? You can't from the physics of the matter. And this gets back to the background interpretive context, the observer relativity, that is the human person where the physics is just... The computer, it's really meaningless without that. And again, that's a very subtle, but I think really powerful point.

It's a point about the human person and I think it's a very traditional and correct insight that there's something about us, whatever you want to call it, other philosophers used to just call it intellect or rationality. And it's also a difference from quality. We haven't even gotten into quality yet. The more traditional philosophers really focused on this sort of determinate semantic content that itself could not be adequately grounded at any materialist conception because no physical particular thing has that type of quality. You take a triangle, it could represent anything, right? It could represent triangularity, it could represent red triangle. Philosopher Ed Feser says it could represent the forgotten Japanese pop band Triangle, for example.

Robert J Marks:

I don't even know what that is.

Pat Flynn:

Neither do I, but it's a fun one to kind of throw out there. I don't know. Maybe you want to talk about that a little bit.

Robert J Marks:

He might've made that up, too.

Pat Flynn:

Yeah, for all I know, right?

Robert J Marks:

Yeah. I-

Pat Flynn:

But since you brought up semantics, I just thought that'd be an important point to put out there. It's a subtle one, but it really pushed my thinking on this matter to say, "Oh, yeah, that's really significant when it comes to the wider philosophical debates of what's going on here."

Robert J Marks:

Well, that is wonderful because we do have these ideas that exist in our brain of idyllic things. I was talking to a philosopher about this incredible number, Chaitin's number. I won't go into the details, but it's an astonishing number. And I said, "Is this platonic?" And he said, "Look, the counting numbers are platonic. One, two, three, four, five. So yeah, this would be that way." But I think it's interesting that you bring up the idea that we have these concepts in our head that we can translate to something like a computer program. And this just goes to reinforcement, that the creativity in computer programs comes from the human being.

We are translating our concepts into the computer.

Pat Flynn:

And we just forget that we've done that at some point, right?

Robert J Marks:

Well, yeah, I think so. I think so. It reminds me of Pointland in Flatland, if you ever read that book. Pointland was the-

Pat Flynn:

Yes.

Robert J Marks:

... yeah, it was this one little point in Flatland and-

Pat Flynn:

Edwin Abbott, right.

Robert J Marks:

... everything that happened. Yeah, Edwin Abbott. Oh, that book by the way, is my top 10. It changed my life. It was really astonishing because after I read it, I thought, "Oh, my gosh, there could be a God, maybe exists a higher dimension." It was really influential on me. But Pointland would sit there and everything that would happen he would go, "Oh, it's so wonderful that I did this." Even if it was talked to externally, it had no concept of anything outside of itself. And I think that that is not exactly the case, but it's very similar to AI programmers that look at computer results and they say, "Ooh, look at what this AI did. It's just incredible," not realizing that, no, that insight, that performance came from them.

And the idea of creativity, and this is important to remember also, is that creativity does not preclude surprising results or results that you would not expect. We get surprising results all the time. We get results that we don't expect. But the question is, can the computer programmer look back and say that this result was beyond the intent or the explanation of the computer programmer? And think about qualia. This is another interesting aspect of artificial intelligence. Qualia, I would say... Well, let's see if you agree. It's kind of a subset of sentience, right?

Pat Flynn:

Right. It's often discussed as a, "what it is likeness."

Robert J Marks:

What it is likeness. Yes. So imagine the following scenario. You are with a man that has been void of the sense of smell and taste since birth. Has no idea. And you bite into a lemon. You bite into the lemon, you feel the crunch of the lemon. You fill all of those little, I don't know what they call them, those little compartments exploding with lemon juice. It hits your tongue. You make this terrible face, and he can see this face. And you experience a qualia, which is the taste of the lemon.

Pat Flynn:

What is likeness of biting into that lemon. Right.

Robert J Marks:

Exactly. So the question is, can you duplicate this experience in this man that has been void of taste and smell since birth? I maintain you can't. You can tell them, "Well, look, lemons are yellow. Okay? They're sour." You can make a face. You can even explain the biology of the taste buds, the chemicals of the acidic lemon juice, all of that stuff, and that helps him understand it. But duplicating that qualia

experience within that man is not possible. Now, if that isn't possible, if you can't explain to a man void of smell and taste since birth, if you can't explain to him, or duplicate, I should say, duplicate that experience, how are you going to be able to duplicate it in a computer program?

Pat Flynn:

How can you make a robot feel itchy?

Robert J Marks:

Yes, yes. Now the interesting thing is that you can...

Pat Flynn:

Cause it to simulate itchy behavior.

Robert J Marks:

Simulate and mimic. They have these artificial tongues, which is just an array of sensors, which look at whatever's put on them, and they kind of get the idea of what the chemical is. And you can train it, when you put lemon juice on it to say, "This is lemon juice, this is sour," and the computer will say, "Ooh, this is sour and this tastes like a lemon." But the question is, in that process, did it duplicate this qualia experience that you have had? And I maintain, no. And I believe that the qualia aspect of sentience is beyond the capability of artificial intelligence or computer programs in general.

Pat Flynn:

It falls into that non-algorithmic bucket, right? Yeah.

Robert J Marks:

Exactly. Exactly.

Pat Flynn:

And it's funny. This is Thomas Nagel's example. It's kind of a gross one, but I think it's funny. You can lick the brain of somebody experiencing the taste of chocolate, and it ain't going to taste like chocolate, right?

Robert J Marks:

I've never heard that. Yeah, that's true.

Pat Flynn:

I might have butchered it. As a matter of fact, philosophy of mind is not my specialty. I peek into it when it becomes relevant for my other research areas, but I wish my friend Jim was here, a co-host, because he does focus on philosophy of mind, so he would be much more adept at this conversation, but maybe we can have a follow-up with him at some point. But, no, I agree with you. So, one of the things that really pushed me was this. Our formal thinking and its determinate exact conceptual content that we have there that really cannot be denied how nothing physical can be like that.

But now we've also focused on, sometimes it's called the qualitative abyss. How could you possibly ever go from something that at bottom is everything that a conscious experience is not? You have these

physical bits that are not intentional, they're not unified, they're not directed about anything. There's no subjectivity there. And we think that if we had enough time or enough ways to swirl this thing together, it somehow will flip over into its almost perfect qualitative opposite, right?

Robert J Marks:

Yes. Yes.

Pat Flynn:

Something that is, has subjectivity experiential states, is unified, is directed or about things. And that seems pretty magical to me. If that happened, I would never think that, "Oh, look, physicalism is true." I would say there has to be more to physical reality than the physicalist's head, otherwise we really are just pulling a rabbit out of a hat type of thing, type of scenario. Nature has to be more like Aristotle thought of it rather than the modern day physicalist thinks about it, if that were even possible to begin with. But a lot of people obviously think, "No, that just isn't possible." So I don't know. Any thoughts or comments on that.

Robert J Marks:

Sure. Yeah, the idea of a monist. A monist is somebody that believes that the brain is a meat computer. They might agree with some of the premises that we made. In fact, I met with a philosopher at University of Wisconsin. We were talking about my visit there. Tononi. Just a brilliant man. And these people, it can't be something spiritual, so there must be a physical explanation. So we have the physical explanation of panpsychism. Are you familiar with that?

Pat Flynn:

Oh, yes. Yeah.

Robert J Marks:

Yeah. Panpsychism is that everything in the world is a little bit conscious, and so therefore you have consciousness, like my mouse here has a little bit of consciousness, and it's just a property of nature. It's like matter and energy and other physical quantities. And they believe that our consciousness comes from that, which, I don't know, I find totally ridiculous. And also there's no background to it. But there's also the belief of some that computers will get to the point someday where there will be this emergence of consciousness. There will be this emergence of the things that we experience.

George Gilder wrote a great book called *Gaming AI*. He called this the rapture of the nerds, okay? That they would get to a point where you would have consciousness which just emerged from all of this complexity. As I understand Tononi's integrated information theory, that's a belief, also. He talks about a measure of the interconnectivity of different things in the brain and believes that once we get to a point, we'll have this emergence.

But I've done some work in emergence, especially like in swarm intelligence and different things, and been a student of the area called artificial life. Really interesting area. And there has been emergence, but it's nothing that is creative. It's not beyond the explanation of the programmer. You see it and you think, "Oh. Well, I programmed it this way. It did this. I wasn't expecting it to do that way, but I can certainly explain what happens." So I don't find that convincing either.

Pat Flynn:

Yeah. And if we're going to have the term emergence be meaningful, we shouldn't be able to just swap in the word magic and lose any sort of explanatory force.

Robert J Marks:

Ha ha ha.

Pat Flynn:

And that's one of my problems with emergence in philosophy, in general philosophy of mind. It's this idea of, part of what motivates panpsychism, and I think this is a good motivation, is they see that emergence isn't enough if we're just using that as a label to cover over a mystery. It's a label that's covering over a mystery. It's not really explaining anything. I think the panpsychist realizes, "Well, it's probably right that we just can't get something from nothing." And if we have a physicalist worldview that it seems to be the case where we don't have the right materials on hand to get this sort of conscious subjectivity. So if we don't want to fall into magical thinking and being a little bit polemical here, maybe we got to have that something already there, right?

Robert J Marks:

Exactly.

Pat Flynn:

You got to have-

Robert J Marks:

And they will go for miles to do that, Pat. It's-

Pat Flynn:

Yeah. And I want to say, I appreciate that motivation. I just don't think you have to go to panpsychism to solve the issue. I think you do have to get rid of physicalism, and you do have to get rid of materialism, but you don't have to get rid... Plus, my issue of panpsychism, if I can rant about it a little bit, is I don't really see how that's compatible with naturalism or physicalism anyways, because that whole worldview is supposed to come through an epistemology that is a scientism, right? That the world is made up of the things that a sort of idealized completed physics and chemistry would tell us exists. And we're not getting panpsychism from that.

Robert J Marks:

Well, no, it doesn't, but I think, Pat, what it's doing is, it's introducing a whole new realm. And this new whole realm is the idea that this psychism is something which is resident in the universe, and it's there, just like matter and energy. So they've added a new axiom, if you will, to naturalism.

Pat Flynn:

Yeah. And again, I'm all fine with coming to see that the universe is a much more special and enchanted place than many modern naturalists think it is. I just think at that point, you're actually just moved away from naturalism, is what I would say at that point. And I would actually say, "You're moving more in what I think is the right direction." I think it is a little bit contrived to keep thinking that this is the naturalist paradigm as we've really understood it, if that makes sense.

Robert J Marks:

Well, here is the current movement. In fact, this was, I believe, introduced by Roger Penrose. I mentioned his book, an Emperor's New Mind. Penrose was a naturalist, and he believed that there were things that the human being did which were beyond the capabilities of algorithms. I picked that up there, and I thought that was just an astonishing conclusion, and the evidence he offered was just compelling. But as a naturalist, he says, "You know, there must be a naturalistic explanation for this." And that's what panpsychism is doing. It's trying to get a naturalistic explanation. But that wasn't Penrose's conclusion. He said that, "I'm thinking around the universe. Where is there anything in the universe in physics that happens that is non-algorithmic?"

And he landed on quantum collapse. When a quantum wave collapses to a deterministic number, he said that is beyond the capability of the description of an algorithm. So he believed that quantum ability, or the quantum effect was what happened when we got consciousness. And he began to work with an anesthesiologist named Hameroff. I forget what the guy's first name is. And they came up with this theory that our consciousness came from quantum effects in our microtubules and our neurons in our brain. And since then we have learned that quantum consciousness is a big area, and this is the last, I believe, unexplored frontier of the attempt of naturalists to explain consciousness.

To my knowledge, there has been no traction experimentally. Everything has been presented in papers and PowerPoint slides, and it would be nice to see some sort of experimental verification other than, "Ooh, quantum happens." I have read papers that say, "Quantum happens." Well, yeah, but that doesn't mean that's the source of quantum mechanics. I just read a paper where there was a quantum interconnection between particles in the brain and particles in the heart, or something like that. I don't know what that means. It could be one of these things which is never proven. It was so outlandish. It's so out there. It might be like string theory. It might be something which is kind of a interesting theory, but which has no traction in reality.

Pat Flynn:

Yes. And again, some of these are really beyond just AI specifically. There are wider philosophical debates in trying to parse out ontological categories. It seems like we just have a category of subjective self-awareness and intentionality and quality and other features of consciousness that really just cannot be reduced to the other category of the physical category, right?

Robert J Marks:

Yeah.

Pat Flynn:

And I think there's definitely a lot going for that. And for people who are interested, I've had conversations around this with many other philosophers of mind on this channel, and actually I think we did a two or part three series on panpsychism as well with Dr. Jim. People can check that out. Critical of it, overall, but acknowledging some of the motivations that we think are good motivations. I don't know if you want to say anything more about those things, Dr. Marks, but I would like to talk a little bit more about what you think, if you saw something from AI that was creative, what would that look like? What would that have to be for you to say, "Oh, wow, this one stumped me," or, "I guess I was wrong on this one."

And then I also, do want to get into... Even if everything that you argue is correct, what concern should we still have about the power of AI? I mean, I just saw a post on Facebook from a professor who just had to deal with the first essay submitted from a student that was AI generated, and the professor's like, "If I

did not know about David Hume as well as I do, I would not have suspected that something fishy was going on about this essay." So there just seems to be a lot of potential negative implications of where AI could go, even granting everything that you argue in your book. I'd like to explore that with you a little bit. I just put two things out there, so please take whichever one you like.

Robert J Marks:

Okay. Well, let's talk about artificial intelligence in general. Is it good or bad? No, it's not good or bad. It's a tool, and it's a tool that can be used for a number of different things. It turns out that the US Patent Office and the US Copyright Office will not generate a patent to a copyright, or will not grant a patent to a copyright, to anything that is generated directly by artificial intelligence. So if you generate an image, a painting, for example, that is generated directly by artificial intelligence, then that can't be copyrighted. In fact, there was just a case in appellate court where a guy was trying to get his AI listed as one of the inventors in a patent, and it was totally overturned. Just-

Pat Flynn:

Interesting.

Robert J Marks:

Yeah. That's very interesting. So, let's see, where was I going with this? There was a guy in Colorado who used AI to generate a painting, and it was at the Colorado State Fair or something, and he won first prize, and it was really a remarkable painting. And so the question was, was he responsible for it? Well, when he was queried about it, he said, "No, I did use AI, but I used it as a tool. I went through and I got a result, and I went back and I changed it a little bit, and then I went back and changed it a little bit." And he went through over 900 iterations. And this is true of any design, by the way. As an engineer, I know that you go through these at... Have you ever heard of Formula 409? Do you know why it's named that? It took Formula 409-

Pat Flynn:

That it took 409 attempts, right? Yeah.

Robert J Marks:

409 attempts. Same thing with WD-40. It took 40 attempts to generate it. So his art was generated with 900 iterations. So the question is, should he be granted a copyright? I don't know if you can patent paintings, but probably not. Should he be granted a copyright as something which was generated by human intellect? I maintain he should, because he used AI as a tool and an iterative tool. And when you use AI as a tool iteratively, yes. All of a sudden you can claim creativity because there's a man in the loop there.

Pat Flynn:

Okay, that answers one of my questions because one of the things I'm interested in for AI is music. I've been a musician all my life. I play guitar and I record guitar tracks, but then I outsource to other musicians and producers. What I would love, and maybe it's already at this point I haven't found it yet, is if I could just take my guitar tracks and have the AI produce everything else around it. That would actually save me a lot of time and I think that'd be pretty cool. But then I'd be like, "Well, could I really claim this is totally my own?" Well, there is a man in the loop there, at least for the guitar parts, right?

Robert J Marks:

Sure, sure. What-

Pat Flynn:

But I wasn't doing the other parts myself, anyways, so whether it's other people... So that's, on the creative aspect, I'm really interested in this because on one thing it could help make certain creative efforts more efficient. At the other hand, I know a lot of musicians that are very worried right now. Very, very worried that, how are they going to be able to keep up in the competition if the AI is able just to crank out so much more music at some point that's just... You can see where certain artists are worried about this.

Robert J Marks:

I will maintain that any music that AI generates will be derivative from its training data. So when you talk about the accompaniment, how do you want the accompaniment? You want the accompany to be in the style of a Wagner, or a Bach, or a-

Pat Flynn:

I'm just going to feel, all of the Van Halen, I can, of course.

Robert J Marks:

Oh, Van Halen? Okay. So you see, there is a predisposition. You've determined what that is. So I maintain that it's probably part of the creative aspect. Here's the interesting thing about the US Copyright Office. You can copyright anything. We talked about this with this attorney, Richard Stevens, and you could just put a bunch of dots on a sheet of paper on musical notations, send it into the copyright office. They don't care. They don't play it. They don't do anything with it. They just say, "Okay, here's a register for the copyright." So if you were to copyright something that was generated by AI, they would give you a copyright. So the only question is, what you do with it? Copyrights and patents, interestingly, only give you the right to sue. And so it doesn't protect you, it just gives you the right to sue other people that use it.

Anyway, I found that fascinating that, yes, even though we have this law that art generated directly by AI is not copyrightable, nobody's going to enforce it. Nobody has the ability to enforce it. It's going to be copyrighted. One of the things that you mentioned about the biggest dangers of artificial intelligence, there's a couple of topics we could talk about here. What are the open problems in AI, and what are the biggest dangers? One of the biggest dangers, I maintain is, number one, having somebody evil getting control of AI. For example, like the Chinese using their facial recognition. I think that that really inhibits the freedom of the Chinese people.

Pat Flynn:

I was going to say, issues of censorship, freedom restriction, propaganda. Those are concerns of mine, for sure.

Robert J Marks:

Yeah. The other big one is unintentional consequences. In other words, and this gets to an area called engineering design ethics. You want to design your system to do what it's supposed to do and not do anything more. And there's classic cases where that hasn't been true of AI. They designed it as best they

could, and then when they deployed it, it did bad things. You hear about self-driving cars, for example, that killed pedestrians. That was an unintended consequence. They overlooked something in the design. And there was an Uber case, I think it was a couple of years ago now that they said that the reason that it didn't respond fast enough was because there was a conflict of the software all trying to do something at the same time and it just didn't work and it slowed it down.

But interesting, the people with self-driving cars and unintended consequences are in the clear. It isn't the fault of the AI. This is also very interesting, that Tesla, I think their self-driving software is called autopilot, was involved in a number of different crashes and fatal collisions, and in no place was Tesla found guilty. Why is that? Because in the engineering design, Tesla realized there might be these unintended consequences. So they told drivers, "You have to still pay attention." In fact, if you're in these self-driving cars and you take your hands off the steering wheel for a while, it says, "Get your hands back on here." And so, when that doesn't happen, it is not the AI's fault because the AI and the instruction of the AI says that you have to keep a human in the loop.

And in that Uber case where a pedestrian was killed, the driver was watching a streaming version of The Voice, or something like that, and was totally distracted from paying attention to the details. And she was instructed not to do that. She was instructed to pay attention. So, unintended consequences is a big effort. In fact, an early example of that was the Russians. They built this system to determine whether or not the United States would do a preemptive military strike of thermonuclear weapons. I think the name of it was called Oku, O-K-U. And the Russian guy that was in charge of that had lights light up and sirens blaring and saying the United States had launched thermonuclear missiles towards Russia.

And first there was one, then there was two, then there was a couple more, and he was ordered to launch a strike back against the United States. His name was Stanislav. And if you're interested, there are YouTube videos where he is interviewed. But he said, "No." He said, "There was something wrong." He said, "If the United States was launching a preemptive strike against the Soviet Union, it wouldn't just be one or two, they would launch like 50 missiles at a time." So he informed the superiors that he didn't think they should launch. They didn't launch, and that saved us from getting into a thermonuclear war.

And that was an unintended consequence of the technology that was developed by the Russians. They later found out that the false alarms came from the sun reflecting off of clouds. So, yeah, they didn't take that into account. And so, the unintended consequences, be it a Uber self-driving car or these thermonuclear strikes or something like that, is something very important and something that we have to pay attention to and the engineering of these artificial intelligence systems.

Pat Flynn:

And that makes sense, even if we don't have to necessarily be too concerned that SkyNet is going to become self-aware, something that may very well be certain, quite still catastrophic, unforeseen consequences that could happen.

Robert J Marks:

Right. Stuff that was in the software, and it doesn't pass the Lovelace Test because in all of these cases, the designers of the AI looked at the software and it did what it was programmed to do. Oku did what it was programmed to do. The Uber self-driving car did what it was programmed to do. So it still doesn't pass the Lovelace Test for any sort of creativity.

Pat Flynn:

Yeah. You know what's funny? It reminds me of, it's been a while since I've seen it, but the old movie Blade Runner. Remember that one?

Robert J Marks:

Yeah. Yeah, I did. Yeah.

Pat Flynn:

And what was the test that they tried to give the AI in the movie?

Robert J Marks:

I don't remember, but I remember them sitting down in a room

Pat Flynn:

Didn't it have something to do with emotion or sympathy or something like that? Which got to one of the points that you brought up earlier, of course. But it didn't seem to just be a... I forget. Now I'm fascinated. I want to go back and watch that movie and see how many predictions did they make?

Robert J Marks:

I got to confess to you, Pat. I remember the movie, but I remember the questions that they asked were questions that I thought, "I don't understand what they're trying to do." I mean, they were really, really strange. Maybe I got to go back and watch it with a little bit more depth of analysis or something.

Pat Flynn:

Yeah, yeah. No, I'm with you. It's been a while, so I probably misremembering as well. Okay, so we've covered a ton of territory here, Dr. Marks. It's been great focusing on just the different categories, ontological categories. There's certain aspects about us that really do seem to be non-algorithmic. There seems like that just creates in principle barriers that AI will not be able to cross. Now we're not saying that it's simulation, right? So there's a distinction here. There might be. It's like an epistemological and metaphysical difference.

Epistemologically, for all we know, it seems to be doing these things, but metaphysically it's really not. And I think everyone assumed, like, "Yeah." An analogy I heard a philosopher use one time is, it's like magic, right? Like, "Wow, a really good magic show really seems like something's happening. But we all know deep down it's not what it appears to be." And the analogy was drawn to AI and its power. It's like, "Wow, it really seems to be like this," but we know deep down upon significant analysis is not actually the case. It really isn't understanding. It really isn't thinking. It isn't really being creative in the sense that you've decided. It's not.

Robert J Marks:

Well, here, Pat, I think is the bottom line. You can't judge a book by its cover. It's an old saying, but you cannot judge AI by its appearance. There is this great website called [thispersondoesnotexist.com](http://thispersondoesnotexist.com). I don't know if you've ever visited it.

Pat Flynn:

I saw it in your book, but I haven't had a chance to go to it yet. Yeah.

Robert J Marks:

It is just fascinating because you go there and you get a picture of a person that looks so real. You hit refresh, it generates another one. And in one of my talks that I gave recently, I put up a picture of eight people and I said, "Here are eight people and four of them are real. Four of them have souls. Four of them have hopes and dreams in life. Four of them have emotions. Four of them were loved by their mothers. The other two are just pixels which are pushed around. Can you differentiate between them?" And the answer is, "No. You couldn't," just by looking at it. Just by looking at the cover of the book. In other words, you have to get deeper inside.

That's the same thing with these chatbots. It's the same thing with a lot of the AI that comes out. New AI comes out and I go, "Wow, maybe they've done something I don't understand." But after I look under the hood and find out what is happening, no, it's still the same old stuff. Not saying that it isn't remarkable, not saying that it isn't astonishing and does wonderful things like generate papers, like you were talking about. But still, it's doing exactly the way the computer programmers were telling it to do. And it is, yeah, it's astonishing, but it's doing what they told them to do. There is no creativity there.

Pat Flynn:

So I would like to finish with your predictions. What is the landscape of artificial intelligence going to look like? How is our world going to change, do you think, over the next 10 to 20 years? Are all of your students going to start handing in plagiarized papers? Or is there going to be some sunshine in here as well?

Robert J Marks:

Well, I think that stuff of the future, I think that there's kind of ongoing problems. One of the things is that artificial intelligence currently has no common sense. There's a story of Fred Flintstone and he got his fingers glued to the bowling ball, and he asked Barney to go out and get a hammer. And Barney went out and got a hammer because they couldn't get Fred's fingers out. And Barney came back and Fred said, "Okay, when I nod my head, hit it. When I nod my head, hit it." So clearly he was referring to the bowling ball, but that vague pronoun was in there, and there's all sorts of common sense problems that really are not solvable yet.

I believe that they can be solved in a very strange way. In some of the places they are solved by GPT-3 because GPT-3 has seen these flubbed headlines and ambiguous things, and figures it out from the context and the syntax of things that it's looked at before. But I think this is a big area of artificial intelligence. One of the interesting things about creativity is that humans experience creativity in what is referred to as flashes of genius. And I don't think you can forecast what artificial intelligence is going to do. I've seen in my career, AI kind of goes along and then boom, there's a great flash of genius that somebody has.

There was back in the '90s something called Arabic propagation. Oh, that's really cool. And then in the last few years, there have been generative adversarial neural networks that generate things such as... Well, it's used in GPT-3. It's used in thispersondoesnotexist.com, and they have come up and you look at them and you go, "Wow, this is creative." I have no idea where this came from, but it really, really works well. Deep learning, convolutional neural networks is another one. So there's going to be stuff that comes on in the future, but interestingly, it's not forecastable because it does take that creative genius in artificial intelligence to go to the next, if I could use the word, quantum step.

Pat Flynn:

Yeah, you certainly can, and I'm glad you did, Dr. Marks. So how about you personally? What projects are you working on? Anything coming down the pipeline? And of course, let's make sure we mention your book again.

Robert J Marks:

Oh, yes.

Pat Flynn:

Non-Computable You, and where people can get all that and all that good stuff.

Robert J Marks:

Yeah. Non-Computable You is available, like everything else in the world, on amazon.com.

Pat Flynn:

Indeed.

Robert J Marks:

And it's available in Kindle, print or audio. And I wrote the book so it can be understood in audio. So that's the good point. All of my other books have equations in it. And I don't know if you've ever listened to an equation being explained in audio. It's...

Pat Flynn:

Not the most exciting.... Yeah.

Robert J Marks:

Oh, it's terrible. It's terrible. So anyway, amazon.com. Non-Computable You: What You Do That Artificial Intelligence Never Will. And it unpacks some of the ideas we've talked about and some additional ideas that AI, for example, will never understand what it's doing. The project that I'm working on now with two guys, Brian Krause and somebody you might know, Angus Menuge, he is a philosopher at Concordia University. He's chairman of the department there. But we're editing a book on the mind-body problem from different aspects. And we're looking at the mind-body problem from different directions, from computer scientists, philosophers, psychologists. And so we have a number of different perspectives on that. And that should be coming out later in 2023.

Pat Flynn:

Oh, well, great. I will keep an eye out for that and have to bug you and maybe some of the other contributors to come back on and talk about that at some point.

Robert J Marks:

Okay. I tell you, I've learned so much. And one of the things about being a professor, you're supposed to re-earn your PhD every seven to 10 years. So I'm re-earning my PhD right now learning about all this philosophical stuff. It's really fun.

Pat Flynn:

Oh, well, God bless you, and you put it to good use with, obviously, your background in computer science where you need to be conversant with these things because you often, sometimes, have the philosophers who have a lot to say on philosophy of mind, but they're just really not conversant with modern AI technology and stuff like that.

Robert J Marks:

Yeah, a lot of them I have a hard time talking to. It's just a different world, but we finally get on a common denominator, so that's good. Yeah.

Pat Flynn:

Now, do you have a website or homepage people can... Or some other place people can look you up?

Robert J Marks:

Yes, in fact, I'm the Director of the Water Bradley Center for Natural and Artificial Intelligence. It's one of the arms of Discovery Institute. Discovery Institute has a number of different centers and the Bradley Center is one of them. The Bradley Center, as the name says, deals with artificial and natural intelligence. So therefore artificial intelligence enters into it. That's been my specialty. We have people writing at our website on different aspects, and these are not journalists. We do have some kind of journalist types, but we also have brain surgeons like Michael Egnor. We have lawyers such as Richard Stephens.

We have psychologists and computer scientists such as Dr. Eric Holloway and Jonathan Bartlett writing on different aspects. And it deals a lot with the natural and artificial intelligence in the latest news. And it adopts my philosophy that humans still are fearfully and wonderfully made, that we are exceptional creations, and it turns out that the evidence is really flowing in. But the website for that is [mindmatters.ai](http://mindmatters.ai). Mindmatters.ai. The politicians say you have to mention your website three times. So let's repeat that. Okay. Mindmatters.ai.

Pat Flynn:

[dot.ai](http://dot.ai). And I'll make it easier because I'll link that and your book in the show notes and I'm going to encourage everybody to pick up a copy. It's a great book. It treats some deep topics, but it's really well written.

Robert J Marks:

Thank you.

Pat Flynn:

Very accessible, entertaining as well. Great voice. So, highly recommend people grab a copy of that. Check out your website. Dr. Marks, thank you so much. It's been an absolute pleasure.

Robert J Marks:

Pat, enjoyed it. God bless.