

# AI and You

Transcript

Guest: Ryan Abbott, part 1

Episode 50

First Aired: Monday, May 31, 2021

Welcome to episode 50! For our half-century episode, I will be talking with Ryan Abbott. He is Professor of Law and Health Sciences at the University of Surrey in the UK, and Adjunct Assistant Professor of Medicine at the David Geffen School of Medicine at UCLA. His research has been featured in the New York Times, Wall Street Journal, and Financial Times, he has been retained as an expert by the United Kingdom Parliament, the European Commission, the World Health Organization, and the World Intellectual Property Organization, and *Managing Intellectual Property* magazine named him as one of the fifty most influential people in intellectual property in 2019.

I'll be talking to him about issues raised in his 2020 book [\*The Reasonable Robot: Artificial Intelligence and the Law\*](#). We've had lawyers on the show before, of course, and we're likely to have many more, because AI is radically transforming law—not just the practice but more importantly, the making, interpretation, and enforcement of law—and the people who do that or are affected by it—and that would be pretty much everyone—are flying to some extent blind, and we want to know as much as we can.

We'll be getting into intellectual property, and whether an AI should be able to own them – and patents, and whether an AI can be an inventor of record, and then we'll be getting into, yes, self-driving cars, the gift that keeps on giving to this show, and even the infamous Trolley Problem – shocking, I know – but some new takes on it and how the law and regulation should adapt, or has or has not adapted. I'm not inclined to explain what the Trolley Problem is again for a while, so if you don't know what that is, listen to episodes 3, 14, 15, 30... 31, 35, or 38. Here we go.

Ryan Abbott, welcome to *AI and You*.

Well, thank you, very excited to be here.

So, you have, in your book, *The Reasonable Robot*, argued for AIs paying taxes, being credited as inventors of patents and being held liable for and punished for crimes. We have got a lot to talk about. This is going to be fun and I hope that we get to nitpick, because I am a nitpicky kind of guy and I figured that's what lawyers do. So, I hope that we can really dissect this in the most pedantic way. Is that all right?

That would be ideal. A few variations on that but we will get into them.

Good I am sure I was painting with too broad a stroke there when I was talking about that, but why don't we start out with you telling us how you got into the field of law, how you got into the field of artificial intelligence, and what brought about the intersection of those in your life?

Sure. Well, how I got into law is a longer story; I think I just like arguing with people when I was a kid. And my dad was a lawyer and a law professor and tried to talk me out of it but at some point, he just gave up. I actually went to medical school and law school at the same time. So, came out not really knowing what I wanted to do but with a whole lot of student debt, and I started working in the area of life sciences intellectual property mainly and I was both teaching patent law and the law of inventorship. At the same time, I was doing work for some life science companies and they were advertising - this was back in 2013, 2014 - that essentially, if you gave them a therapeutic target, they would have an artificial intelligence go through a large library of potential drug candidates and pick out one for you and give you a bunch of data suggesting it might work. And I was looking at that, and thinking, "Gee that's interesting because when a person does that, we file a patent on that and they become an inventor." So, what happens if you have a machine do it and I wonder if anyone's thought about that before? Well, it turned out people have been writing about it since the 60s, so I was hardly the first person to think about it but people were also kind of writing like, well if you had a machine make this sort of thing, you wouldn't really need a patent on it because machines don't care about incentives. And while that's certainly right, the people who need a patent on it are the drug companies that are investing in building complex AI and using it and thus then have to kind of do follow-on clinical trials on a drug. So, I sort of thought, well actually you still need a patent just as much; you just wouldn't have a traditional human inventor maybe. So, that was kind of my first time I looked at AI doing human sorts of things but maybe having very different legal frameworks in place for the same sorts of behavior and thinking that maybe that didn't work so well.

I am going to be getting into this in a lot more nitpicky detail, but I think that the thing that's occurring to me at the moment is that when a lawyer gets into writing this kind of book, which goes into case precedent and a lot of detail that takes a tremendous amount of work and research; when someone in your field puts that kind of investment into this, it indicates that the field has reached a certain level of maturity, that it is something that's about to become meaningful and be worth real value to people and corporations. Is that how you see it?

Absolutely. I think at this point, you know, as you have talked about in some of your work, there's a wide disagreement on what the future of AI is going to be, whether it is a utopia or a dystopia, or even whether it is going to have dramatic economic consequences or not. In one view of the future - and this is kind of where my tax work comes in - AI has put us all out of work. Maybe that's good, maybe that's bad; and then there is a version in which AI kind of doesn't do that much and actually people just find new sorts of things to do. But certainly I think in the not too distant future AI is going to have a dramatic impact in at least some areas. One place I think this will happen is in intellectual property; less with patents but more with copyright. So, if you go listen to something like Open AI's jukebox initiative, you can hear an AI that has made like a passable Elvis Presley, Katy Perry and Frank Sinatra song and they are not *that* good but they are not *awful* either. They can write short articles that people cannot tell or written by a machine and I think as they keep getting better at this in the next five to ten years, AI may generate a tremendous amount of creative output that has some value even if it's mediocre and that AI will cost almost nothing to do this.

So, broadly, what is the intent of your book here when you are making arguments for treating AIs in categories that we have previously only associated with humans? What's the umbrella term for all of that and what are you trying to achieve with that kind of equivalence?

Right. Well, I wouldn't say it is treating AI like a person because an AI isn't like a person, I would say it is treating AI behavior the same way we treat human behavior and it is a subtle difference but one that I think is kind of legally important. So, for example, let's say that instead of hiring someone from my company to do new graphic designs, I licensed an artificial intelligence that could make new graphic designs and so essentially, you have an AI and a person doing the same sort of functional behavior for a company and granted there are some ways in which it might differ but from a business case, the company might say, "These are more or less exchangeable, do we want a person or a machine?" I am not in favor of AI paying taxes *per se* but one of the interesting things that happens when a company automates like this is they save money on taxes because in the US, you have to make payroll contributions for an employee - or in the UK, you have to make national insurance contributions - so you are paying the government to employ a person. You don't pay the government when you employ an AI and so our current tax policies artificially encourage companies to automate. There's other ways that they save on taxes by doing that, for instance, by accelerating certain deductions associated with AI but not people, or avoiding indirect tax consequences that are more nitpicky than most people want to talk about but they're in the book. The gist of this is that the law in the tax area discriminates between having a person do something, having a machine do something, and it has unintended and negative consequences. And if the law were more neutral to AI and human behavior, I think we would often better realize the law's underlying goals and I call that a principle of AI legal neutrality.

Yes, so you have used that term "legal neutrality" for AIs in the book and I am trying to understand whether there's a line being crossed here. You talk about AIs as an entity as though that has a legal meaning and I imagine you don't intend that, but I want to see where we have crossed the line from tool to something that has some kind of implied agency. If I use a hammer to build a house, the hammer isn't the creator of the house but in your language it sounds as though if I have a robot go and put up drywall - and those exist - it has a different kind of ownership in this situation. Are we getting closer to the mark?

Well, I think these are indeed very interesting and very complicated discussions to be having. The intellectual property field is a good way in which that sort of thing is illustrated. So, for example, right, I use Microsoft Word to help me write things. It makes suggestions. PowerPoint is getting pretty good at helping me put together slide decks; I wouldn't think of it as an author, I wouldn't think of it as a joint contributor to a work. On the other hand, there are commercially available programs right now where you can put in a sentence or a couple of themes and it will just generate an article for you. Functionally, that AI is doing the sort of thing that a human author used to do. Now, *legally*, could an AI be an author? Legally, could an AI have any sort of ownership? The first question is a little more complicated than the second. Under current legal regimes an AI doesn't have legal personality, it couldn't own anything, it *is* something that is owned. So, (a) generally right now the answer is just No and (b) also it *should* be No because it

wouldn't make a whole lot of sense. An AI wouldn't care if it gets copyright in something, an AI couldn't license something that was copywritten without some very convoluted legal kind of guardianship regime. An AI wouldn't enforce its intellectual property, there's really no reason for an AI to own something. On the other hand, traditionally in the US to have a piece of work that has copyright protection, you have to have an author, right? So, what do you do if the functional author is a machine and not a human being? Who would you list as an author if anyone, could that sort of thing even be protected? Well, the United Kingdom back in 1988 passed the first law that said, if an AI makes something that can get copyright, you don't have a traditional human author. We are going to legally fictionalize the author as the producer of the work, like a film producer and you will get a shortened term of statutory protection. So, we kind of have this specialized protection regime for that and a legal fiction. The US, on the other hand, says no, no, if you don't have a traditional human author, you cannot get protection at all. So, if I make an AI that can make really good Katy Perry songs, maybe even better than Katy Perry Katy Perry songs, there's no economic value to me doing that at least from a copyright perspective. And you see that's what I think is a bad outcome; that we want super Katy Perry AIs, right, we need them as a society and the way to encourage that, or the way we traditionally encourage that sort of activity is through copyright protection. So, I would suggest that on something like that made that the AI might be listed as the author on it not to provide a machine with rights or credit - it cares of neither - but to state this is an AI-generated work so the public knows how the work was done, and to keep people from claiming credit they don't deserve. So, let's say I made an AI that made a thousand Katy Perry songs, and Katy Perry owned it. Would we want her to say, I made all this music? If she did not really make it, if really a machine made it but she owned the machine, I think that changes what it means to be a human author in a negative sort of way.

So, is this about the amount of effort involved in the creation of the work, like if I sit down at a word processor and I am typing every word and it suggests things and corrects my spelling as I go along then I am the author of the work. But if I push a button and something like GPT-3 makes up a three-page story, then I am not entitled to claim authorship, is that both sides of a line you are drawing?

Yeah, I think that sounds about right. There is definitely I am kind of the functional, who-did-what-when part, I mean that's a critical part of these individual analysis, right? I mean, I think it is very clear if you write something in Microsoft Word, the machine has done very little even if it corrects some stuff, if you put three sentences in GPT-3 and it generates a three-page article for you, I think the opposite's the case, the machine's done the work. In between, yes, there is a fine line. But on the other hand, this is something we do with people; so if you had a co-author on your book or maybe you had people copying your book, do they become co-authors, is someone a co-author of a chapter? The law has already developed rules for dealing with when that line is crossed, and dealing with edge cases and I think more or less those rules apply to machine contributions.

Incidentally, I will point out that someone who was on an earlier episode of this podcast, Roman Yampolskiy, Professor of computer science at the University of Louisville, Kentucky, has credited a co-author on at least one of his papers as "MS Spell Checker" and got it published as

such. I think he's trying to test some of the same limits that you are, which brings me to: So what's the utility of drawing these distinctions at this point? Because is there a case before a court somewhere that this would inform, are you anticipating that? Are you carving out a space in the thought platform here? When will this become - it doesn't seem like this is real yet. It seems like this is very hypothetical - and I totally dig hypothetical - but when and how is it going to become real?

Well, it is definitely going to be, I mean it is definitely already real. It is interesting that your former colleague or that your former host did that. I was trying to think of who did it earlier but I think in the '50s, I am going to misremember the names of kind of the AI pioneers who had, I think an AI called Logic Theorist. I was going to say, I hope we are not recording this but we are, I cannot remember the name of the AI or the researchers but they are in their book and just as they had it solved this math equation and they submitted the solution to a journal and tried to list it as a co-author and the journal said, no forget it. That is a little bit vanity. Of course, the case I am involved with is what I like talking about. So, I will mention it and maybe I will talk about the Siemens case too. The world intellectual property organization, which is the UN agency that kind of is responsible for IP among others has been having a series of international conversations on AI and IP. And in 2019, Siemens presented a case study where they had an AI that made a new industrial component. They wanted to file a patent for it and couldn't because they said, they asked their engineers to be inventors and the engineers basically all said, well we told the machine what we wanted, it spit out this really interesting design, we all looked at it and thought wow that's cool and great but we did not invent this thing and we are not going to put our name on it. Because of that, they were unable to file for a patent, and it isn't just vanity on the part of the Siemens engineers either because putting yourself inaccurately on a patent in the United States is a criminal offense. I have been spearheading a series of legal test cases with a group of international patent attorneys from an AI called Dabus, which was designed by Dr. Stephen Thaler in the United States. And Dabus generated two inventions without a traditional human inventor. One for a flashing light that could attract human attention in an emergency, and one for a beverage container based on fractal geometry. We filed these in the UK and in Europe and they both received preliminary indications, they were patentable and then when we said well actually there was not a human inventor, both of these got rejected. And those applications have been rejected in the US, Australia, Germany, the UK, and Europe and all of these are under appeal right now and courts are looking at, well can you have an invention without a human inventor? Who or what would you list as an inventor and who would own the thing and these are really unanswered questions that patent law had not grappled with when it was being written, but now in light of technological advances has to deal with. And I think the copyright side is much further along than this. There has not been a lot of litigation about it but also AI has kind of only recently gotten to the point where it has been writing books for decades. They have just been awful and no one would read them except for novelty but they are not so awful anymore. And in 10 years, they might be better than my books.

And I have heard some of the music that's been composed by AI and some of it is better than what I have heard many humans create. So, I agree that there's value being generated there

and you say also in your book, you say, more importantly sometimes people will need to be treated more like AI, what do you mean by that?

Sure. Well, to give you another example of where this has already come in, and to illustrate that point: liability for harms caused by algorithms. So, AI is doing a lot more and it is doing human-like sorts of things and human-like sorts of things cause harm, whether you have an AI working as a robotic tool or helping to diagnose people based on imaging or what we are soon to see on our streets, everywhere we look are self-driving cars. Self-driving cars are interesting because there you have an AI doing something that is really very functionally interchangeable with a person, right? So, I may summon an Uber and not know if I am going to get a human driver or an AI driver. Also, interestingly here is that an accident caused by a human driver and an AI are subject to very different standards, both legally and otherwise. For example, there has been at least one fatality caused by a self-driving car, which was in Uber in Arizona and that fatality got a tremendous amount of press coverage. 30,000 people a year are killed by human drivers in the United States and this gets comparatively little. 94% of those accidents are caused by human error and so it is actually likely whether or not it is the case now that self-driving cars will not be perfect; they will hurt people, but they are likely to be substantially safer than human drivers. And so putting them on the road is likely to result in aggregate safety gains, even if they still cause accidents, and yet we hold them to a much higher standard of legal liability. So when a person causes an accident, we look at that in negligence and say, "Would a reasonable human driver have caused that accident, if yes then you are not liable for it, if no then you are." But with machines, we hold them to strict liability standards. We ask - it is more complicated than this but we basically asked - whether the machine caused an accident and if there was a defect and then it is liable for it. So, this discourages the use of machines and driving because we hold them to higher liability standards and there is more liability for using them than for using a human driver. And that may discourage the use of technologies that actually improve safety. Now, in the long term, because they are going to be so much safer, that may not make a huge difference because even with higher liability, there will be so many fewer accidents, it won't matter as much to manufacturers. But then kind of something interesting happens where it is now not a matter of is it safe to introduce a self-driving car, it's more, is it safe to have human drivers? If self-driving cars are almost perfectly safe, do we want humans on the road and Elon Musk has said, "No, we want to ban them." There are some good reasons not to ban people though. We don't want to limit human autonomy and freedom. On the other hand, you are not only risking your own safety getting behind a wheel, you are risking mine as a pedestrian. So, one of the things the book says is, well when we have self-driving cars that are much safer than people, maybe we want to hold a person to the standard of a self-driving car. In other words, if I cause an accident driving and I could have been in a self-driving car, would a self-driving car have caused that accident? And if yes, I wouldn't be liable, and if no, I would be. And that wouldn't keep me off the road but it meant that if I chose not to automate and automating would have been safer then I will be financially liable for the harms I cause.

I see an analogy with imagining you are on a commercial airline flight and the pilot announces, hey everyone, I just want you to know that I am turning off the autopilot because that's how I roll and it is more fun this way. I think we would consider that reckless. Now, I think the

question about whether or when AI will be a safer driver, is still up in the air but that what will approach sooner than AI being universally safe, is a situation where on average AI is safer. Say on average, they are ten times safer, ten times as few accidents with self-driving vehicles. But there will be situations that they get into where they cause fatal accidents that humans almost certainly would have avoided and would be held culpable for in similar circumstances and that puts us a kind of a version of the Trolley Problem, which listeners of this show have heard about... several times... by now but that we have to decide, are we going to deploy something which will on average cause fewer deaths? If it results in deaths that wouldn't otherwise have occurred, where does that fit into the law, what's going to happen? Who stands to benefit, or how do the different parties involved there exercise that kind of case?

Now those are great questions. Indeed, I think it gets simpler when you have black and white answers like cars are dramatically always safer than a human driver and for a long time, it may be that cars are safer in some circumstances and people are in others. And I think the near-term future is one in which there is some optimal mix of human and AI activity and that's how we get our best outcomes. So, maybe self-driving cars don't drive well in snowy conditions and people know to turn their car off and there are all sorts of challenges with that, some of the Tesla fatalities whether or not they were caused by the car, Tesla expects you to have your hands on the wheel with the cars driving but of course everyone realizes that that just will not hold up in a five-hour drive, someone's going to lose attention after they have not had to do anything that whole time. People may get de-skilled by letting cars take things over. I think these are really very challenging issues that have to be resolved on a case-by-case basis by making decisions carefully weighing the costs and benefits of that, and there are a number of parties who will before the fact and after the fact be weighing in on whether this is something we want to do; regulators, individual manufacturers, individual consumers, lawyers after the fact suing over accidents caused by cars, they are really complicated answers that I think one has to look very specifically at. The Trolley Problem is interesting also in the context of cars, I guess we will not describe it for the audience because they have heard it enough but I think a couple things about self-driving cars are interesting that are not so obviously self-evident – that's kind of redundant - but one of which is human drivers make trolley problem decisions occasionally but have to make them in split seconds. Rarely is someone driving down the street and thinking, "Gee should I hit the two people or the one person or the old person and the young person." We are pretty clear better to hit one person than two, all things being equal. But generally, someone isn't kind of able to rationally think through, should I get the older the young person and if they are, they could probably just slam on the brakes. Self-driving cars, I think, are good because in some sense they are going to remove the need for dealing with the Trolley Problem because they will have superhuman reaction times and rather than do I hit the old or young person, let me choose option C and stop the car faster than a human could. But the other challenge is that while we don't demand that people have thought these things through and they don't all have clear ethical or legal solutions to them, car programmers do have to think these problems through now because they have the chance to decide what the car will do in advance, or decide that they are going to either abrogate some responsibility to say a driver, to have a driver's ethical preference for these sorts of things, or try and fudge things at the programming level so they can avoid trying to

distinguish between these like, for example, not collecting the age of pedestrians and so only having a chance of hitting one of two pedestrians and that being equal, hit one with less force if possible. So, I think the long rambling answer was, these are very interesting ethical and legal issues and one really has to decide them on a case-by-case basis.

Isn't that line of reasoning - which is not uncommon about the pre-programming of the AI, it has been pre-programmed to deal with these situations - isn't that at variance, though, with your earlier examples like say, of the Siemens patent where the engineers said, well it isn't ours, the AI was responsible for that. And the programming of a self-driving car involves a lot of training certainly, but not of every conceivable situation that it will face, there will be some where it will have to make up the answer on the fly, and those are likely to be the ones where those accidents occur. I would like to know but I rather doubt that the people who train self-driving vehicles actually do real-life versions of the Trolley Problem just to put that case into the programming.

Well, the good people do, right? So, one of the differences with self-driving cars and an AI that is generating new industrial designs for human teams to evaluate is that there is a lot more risk with the self-driving car case than with spitting out an industrial design nobody likes. If you are a company working commercially in this space, these are issues you are aware of and hopefully have strong governance mechanisms involved where you have both C-suite level engagement and upper management engagement and multi-disciplinary teams looking at these challenges with programmers and deciding how you are going to respond to these. That doesn't mean every single instance is done, but these should involve teams, and some of them definitely do involve teams, that are looking at, you know, if for the Trolley Problem being something like, do I hit more fewer people, I mean the answer there is you hit fewer people and no one's really debating that. The Trolley Problem gets more complex where there's action or inaction causing problems, or demographics of users. So maybe teams are deciding, do we even want to collect these demographics, do we want the car to be capable of making a decision of who to hit based on demographics, and probably not; and then to the extent, what sort of factors impact the car's decision-making ability? How do we want the car to respond based on those factors, what factors do we want it to have and so forth? So, I think that people who were doing this right are looking at these things and thinking them through and we wouldn't want an unpredictable neural network out on the street kind of making decisions on the fly for the first time in that context.

Well, there's a spectrum of these situations. So, what we have at the moment is that we can train AIs on situations like, I am following a truck and a motorcycle cuts in front and it will work regardless of the color of the motorcycle, size of the motorcycle, whether there's someone riding on the pillion, whether it is a panel truck or an RV, that kind of thing. It can generalize that way, but if it has not been trained on a situation of say a child riding a tricycle around a roundabout the wrong direction, it has never seen that before; it may have no idea, and there are so many of those kind of long tail situations in driving in first and last mile street environments, that the behavior is likely to be unpredictable and so it may fall back on general programming of "avoid killing people, just stop wherever possible," but I find it hard to believe

that they are programming in situations which are engineered as, “choose this person over this one in such a situation.” For one thing it seems to open up to embarrassing revelations in discovery.

Well, I mean indeed these are very complex issues. Let me talk about both of the ones you did. So, for the motorcycle driver, for instance, a car may have a choice of hitting - to the extent that the Trolley Problem is something that drivers confront that often and that the car doesn't have option C, which is stop the car, don't hit anyone; but if a car is forced to hit another car or a motorcycle, one of the things that a car might take into account is which one of those is better to hit. Hitting a motorcycle is probably better for the passengers in the car, probably not so good for the object being hit. Another thing is it might take into account whether the driver of a motorcycle is wearing a helmet, I mean that might be something a human driver takes into account. These are the sorts of things that programming teams should know to be looking at and addressing and deciding, right? Even making a conscious decision “We are not going to collect information on whether a motorcycle driver is using a helmet, we will not allow that to influence a car's decision making.” You know, left to its own devices, a car may generate some kind of rule for dealing with that. The kid on the tricycle also is something that, I think is not such an unknown sort of thing that a team couldn't possibly design for it. Now, as a practical matter, I do agree, things will come up to cars that it hasn't been programmed to deal with; but the machine should still have a framework for evaluating new situations that address which factors it takes into account and what decisions it makes and even more objectively technical things like, we don't care if the person on the motorcycle is using a helmet or not, I mean that is even there a value decision and something that a computer scientist working in isolation shouldn't be deciding on his or her own, right? It should be having kind of corporate and multi-disciplinary input.

Sure, and one of the things that surprised me in this space of self-driving vehicles, when I got a Tesla, I was shocked that I could just tell it, drive yourself, in areas where there was no regulation whereas the vehicle code is full of regulation, I cannot turn on a turn signal at the wrong place without potentially getting a ticket. It is that specified. The California vehicle code specifies the range of wavelengths of light that turn signals and stop lights, brake lights, can have. And yet it is so silent, relatively speaking, about autonomous vehicles. I was expecting that this wouldn't happen until there was regulation saying, you are allowed to engage them in these streets, in these lanes, under these conditions, that you have to have these licenses. But we don't even have the indication on the outside the vehicle that it is under autonomous control, except where it has been chosen to put there. It seems as though these things have pardon the pun, driven away ahead of legislation, and so is the law behind in this area? Does it have to evolve to catch up?

Oh, another great point. It has long been the saying that the law catches up with technology, right? And sometimes maybe that's for the best and sometimes it is probably not and I think this is probably one of the examples where it it's not so good because AI has both the potential to generate massive social benefits but also to cause very substantial harms and it has already been used to do that, whether it is killing pedestrians driving a car, causing flash crashes in the stock

market, causing political manipulation, the risks of AI are substantial. And generally the way that we as a society try to maximize social benefits and reduce risks from harm is through our framework of laws and regulations. And this is one reason why the law shouldn't just sort of wait to see what the market develops from things and move ahead. Which is not to say the law is the only way we get those things. Corporate responsibility, for example, plays a big role in this industry standards. The law says kind of a floor for we have decided we definitely don't want people doing this sort of thing and there are still higher ethical and corporate bars for activity. Which isn't to say that AI is unregulated either, I mean AI is regulated just by regulations that are from decades ago which did not have AI necessarily in mind, and so this book is one effort to say your existing regulations are not getting the most out of AI we can be getting, we should really be seriously rethinking some of these things as a society and one of the principles that will help us come up with better regulations is by focusing on behavior regardless of the nature of the actor and in regulating behavior in ways that get us the benefits we want out of the law. I do think this is something particularly in recent years that regulators have been very attuned to, especially in the self-driving car context. The UK, Europe, the US have all had regulatory and legislative efforts related to artificial intelligence; I think the European commission will be coming out with more detailed regulations soon, and already has in a few areas. So, COVID kind of knocked it off track, but I think there is a growing understanding that the law shouldn't just let the market do whatever it comes up with.

That's the end of part 1 of the interview. By the way, Ryan mentioned Logic Theorist there, so I'll give you some context about that. That was a program written right at the dawn of AI, in other words in 1956, and was presented at the summer conference at Dartmouth College where the term AI was born. It was a program that was able to prove 38 of the first 52 theorems in chapter 2 of Whitehead and Russell's *Principia Mathematica*, one of which was actually better than the human solution. Not bad for whatever computing hardware, let alone software, they had in 1956.

In today's news ripped from the headlines about AI, Amit Gupta and James Yu created a new app called sudowrite. I would hazard a guess that the 'sudo' part comes from the Unix command for using root privileges, because sudowrite uses GPT-3 under the covers to fill in writing prompts. It was the subject of a [New Yorker article](#) whose author had a hankering for Samuel Taylor Coleridge, the poet famous for the poem *Kubla Khan*, which starts out, "In Xanadu did Kubla Khan / A stately pleasure-dome decree: / Where Alph, the sacred river, ran / Through caverns measureless to man / Down to a sunless sea." Unfortunately, while Coleridge allegedly had the next three hundred lines in his head as a result of an opium dream he had just awakened from, he didn't get much more down on paper before the doorbell rang, and by the time he was done with the infamous "Visitor from Porlock," who was probably selling vinyl siding or car insurance, he had forgotten it. We've all been there, right? Some of you software developers know what it's like to have a whole program in your head only to have it evaporate when the guy in the next cubicle asks if you've got any extra coffee creamer.

Well, the *New Yorker* author, Stephen Marche, curious what might have come next, fed the beginning of *Kubla Khan* to Sudowrite, and it went on with: "The old Moon shook her silver tresses / The new Moon bow'd her golden horns: / While all the stars that round her burn'd, / Bow'd to the ground and based their fires." And it got the poet-y-like apostrophes and indentation, and wrote a lot more in the same

vein. Marche said, “I find this beautiful, memorable. If you told me that Coleridge wrote it, I would believe you,” and I agree.

There are many inquiries prompted by this. (By the way, if you’re expecting a punchline that this epilogue was written by Sudowrite, I’ll head that one off right now by saying that it isn’t. But if I ever do something like that, I’ll tell you. Eventually.) The usual debate is about whether computers can be truly creative and arguing whether that kind of example is really beautiful. That’s not a productive question to ask. Yes, it’s imitating the style of its writing prompt, but it didn’t need much to create new text. Whereas even Coleridge was copying some of what he wrote from the book he was reading when he fell asleep, and yet the phrases that Sudowrite came up with cannot be found on the Internet, part of which was its training corpus. In other words, it wasn’t cheating by regurgitating some other poet’s attempt at ripping off Coleridge. Another interesting question would be what this means for students. If all a high school student needs to do with an essay is write the first two sentences, and Sudowrite can complete the assignment in the student’s style without triggering any plagiarism detectors, the future of homework could be radically different. Not to mention the future of a lot of writing assignments that are not demanding. We already have AIs that take financial or sports game results and weave some prose around those statistics to engage the reader in a story that has to be written a hundred times a day, hitherto by some poor writer who has exhausted every possible way of saying “This team slash company was utterly defeated.” And then there’s the question of who owns the content it creates, and you’ve already heard me discuss that with Ryan.

A quick plug, for those of you listening to this before June 3, 2021, when I will be appearing live on the [Strategy and Leadership Podcast](#), by Anthony Taylor, at 9am Pacific time, link in the show notes and transcript.

Next week, I’ll be concluding the interview with Ryan Abbott, when we’ll be talking more about self-driving cars, but now in the context of the matter of intent, and liability and punishment for AI infractions, and rationales for changing our taxation system for AI-generated wealth.

That’s next week on *AI and You*.

Until then, remember: no matter how much computers learn how to do, it’s how we come together as *humans* that matters.

<http://aiandyou.net>