

AI and You

Transcript

Guest: Tom White

Episode 103

First Aired: Monday, June 6, 2022

Hello, and welcome to episode 103! My guest today is Tom White, an artist, and he investigates what he calls the *Algorithmic Gaze*: how machines see, know, and articulate the world. How does he do that, you may ask, and what does it look like? He uses neural networks to produce abstract ink prints that AI engines decide are the very essence of some concept that he has picked, like “Rabbit”, “Banana”, or “Dalmatian”. But to you or me, they look like someone made puddles of one or two colors and then scratched black lines of different thicknesses across them at angles with a black marker pen. They’re very stylized, very recognizable, very minimal. And they are very abstract. Few of them look like anything in particular, but sometimes, if you know what the image recognition AI thinks it is, you can see the general idea.

But the AI is dead certain what it is, more so than for just about any real image you could give it. This is the special secret of his work, which he has exhibited all over the world – Taiwan, Zurich, New York, Sundance, Montreal, India, Japan, Italy, Monaco, and more, in real prints hanging on real gallery walls. If you take a picture of one of these abstract prints with your smart phone, say you photograph the one of a banana, which looks like a yellow bean bag with graffiti, it will likely be classified as food.

And Tom has done some more *interesting* work than bananas, as we’ll get to in the interview. He is also a lecturer teaching computational design and creative AI. Here’s Tom White, calling in from New Zealand.

Tom White, welcome to AI and You.

Hi, thanks for having me.

So, how did you first get into this idea of art interacting with artificial intelligence?

It has two phases really. I’m from a kind of an original generation, or an older generation of generative artists back in the late ’90s and so at the time, there was a group of artists that were doing artificial intelligence-based work, people like Carl Sims, but it was using a different technology. Also in that era was Harold Cohen. That was my interest, was sort of combining computer perception and computer abilities with art. But then with the AI winter, I kind of did other things for a while. But then when AI got interesting again, about 10 years ago, I started looking into the latest libraries and various capabilities and I started building kind of a new series of work based on computer perception, because it had gotten so much better in the 10 or 15 years since then.

What, from your perspective, changed in AI a decade ago, qualitatively?

So, when I was in graduate school, they did have computer vision systems, but that was in the era of good old-fashioned AI. So, that was, if you wanted to build a system, you might build something called an expert system, and you would kind of break something down logically and you would say, well, I want to recognize a face, that face is going to have two eyes, it's going to have a nose, I'd better build something that can find the eyes and build something that can find the nose and you kind of deconstruct the problem. What changed about 10 years ago was that machine learning really took off and using the technology of neural nets. Now machine learning is different. It's a subset of artificial intelligence in that it builds its own representations, so that there's less bias and there's less kind of preparation. You build a data set and the data set is a bunch of faces you want to recognize, and the computer, the algorithm is what determines what's important to look at. So, you don't really feature engineer, you don't say I need to find two eyes, I need to find a nose. It's really the algorithm internally that decides what the best representations are to get the job done.

That good old fashioned artificial intelligence of having to spell out what a nose is and it sounds like the antithesis of art. It just sounds so left brain.

It's a good point, I think that there's probably my hero of that era is Harold Cohen. So, Harold Cohen was an abstract painter and he went into artificial intelligence in the 70s and he essentially built a system that could do a lot of the drawing and painting and coloring for him. So, he had a set of principles that he operated on, and then he just folded those into an algorithm and the algorithm could generate scenes of plant scenes of people. And then he even had his own painting machine that filled it in, I guess I'm kind of pushing back on the idea that it's not artificial intelligence; I think artificial intelligence is kind of defined by the era it lives in. So, at the time, this was a very advanced capability, this idea that the algorithm could kind of endlessly generate variations of people and plants. But you're right: now, looking back on it. Now, this is not the way we deconstruct the it's not the way that artificial intelligence kind of best practice deconstructs the problem, what they use is what's called end to end learning, where we really just feed data into the into the algorithm in the neural net is what makes the powerful is the hidden layers and the hidden layers are where the representations of what the algorithms think are important to get the job done, or formed.

So, what are you doing, what's your goal, from a high-level perspective? Because you're creating art, but with what aim in mind? To make us think, to show how AI can be fooled - fill in the blank.

Yeah, so I think there's different ways of approaching my goals. The most approachable is just to do a thought experiment where you imagine we encounter under an alien culture, so if they were aliens that came down to earth, and they had different bodies than us, and they had different perceptual qualities they did, maybe their eyes didn't work the same, or they experienced reality in a different way, it would really be interesting to see what kind of art these aliens generated as a way of understanding them. To me, the machine learning systems are kind of an alien culture, they have ways of thinking, they have ways of seeing and understanding, and they're built in our image, but they're not exactly how we work and so I think there's something really interesting in making art, leveraging this kind of foreign capabilities and foreign kind of supernatural powers

that the AI has. And so my goal is to make representational art. So, I'm not making abstract a lot of generative art and a lot of computer art is I don't want to belittle it, but it looks it's kind of like spirograph art, where you it's about patterns in sort of the algorithm itself, and how you can do something with repetition. My interest is more representational, which is, what makes a bee a bee, or a flower or a flower? And how is the how is the machine learning version of those concepts different than our own? If you have a machine that knows a crab from a snail? Is it using the same internal logic to do that, and what does it look like if you ask it to build art using its own representations?

And so, you're talking about how AI sees the world, we know that it sees it differently from the way that we do. You refer to something called the *algorithmic gaze*. Can you tell us what that is?

Sure, I think that increasingly, a good analogy to this might be text prediction. So, 10 years ago, 15 years ago, the idea that that an algorithm would be able to predict what you're going to say, was kind of a foreign idea. But now that we have autocomplete in Google, and certainly when you're texting someone, you have the algorithm always trying to guess and second guess what you're doing. And I think more and more this is going to be happening with vision systems. So, if you have a self-driving car, if you have a security system, it's going to be helping you and it's going to be kind of anticipating what it sees. And just like when you're using text prediction it makes a bunch of kind of boneheaded mistakes, and you're trying to send a message and most of the time, it's helpful and most sometimes it's not, I think that the vision systems in the same way, kind of have their own internal logic and so the algorithmic days refers to this idea that these systems have their own way of seeing the world, which in some ways is very aligned to ours, and sometimes it's different. And what I'm trying to do is surface this as natively as I can. So, by using machine learning in my process, then whatever unique characteristics or capabilities these systems have kind of bubbles up in the art itself. And a lot of times, this surprises me. So, I don't really go into this thinking, I'm going to show how the algorithm does A, B, or C, I just kind of make the art using the algorithm and then a lot of the qualities of machine learning kind of kind of surface naturally in that way.

And so, there's this idea in philosophy that we don't know what it's like to be an x, where X could be another person, like they say, that's red and we look at that, and we say that to red to we don't know what it's like for them to experience red. So, it sounds like you're trying to show us what it's like to look through the AI's eyes at the world.

That's a very good analogy. So, I think it was Thomas Nagel, that published the essay, what is it like to be a bat, and he talks about how, what can we even imagine? What a bat who doesn't have vision and traverses the world with sonar? Like, how can we even talk about the life experience of a bat? And I think that's a really good analogy, because where I'm trying to do something very similar, but it's what is the inner life of these algorithms, which we're building in our image, and we're building them we're building them with to do certain tasks, but what is their kind of inner perspective? And I should say that this isn't a this is a very much a moving target. So, if we look at what my art was doing three years ago, and what more it was doing maybe a year ago, there's actually going to be different answers because machine learning is

such a rapidly evolving field that the body of work is kind of trying tracking what the what the latest fads in machine learning are.

Now, to get to how the art is produced. Are you drawing this out? Or do you get a computer to do it, what's that process like?

So, I've always been interested in this work in generating physical work, or that's been the goal. The physical work is meant to be perceived by the computer. But it's worth pointing that out at the beginning, because there's no one right way to view a physical work. So, if you've ever done any photography, you know that different lighting conditions, different angles, make things different. So, the, the starting point for me is to make something physical and part of this too, is that I said, I've been I've been doing computer art for many years and I think what artificial intelligence is doing more than anything, it's kind of removing the digital-analogue barrier. So, things kind of hop across the barrier now, with tremendous ease. So, you can make physical work, you take your camera out, you take a photo of it, it goes to Instagram, Instagram's algorithms are instantly looking at it and sorting it for you. So, I think in the past where I've done sort of installation work, it's more natural now to do physical work, because it can go digital right away. It's a little bit of an aside there. Can you repeat the question for me? I got a little sidetracked.

What's it like? Show us through your eyes what it's like to create something? How do you get an idea? How do you go about it? How does it end up coming out?

Sure. So, one goal might be - when I was starting off, the it was in the era where most machine learning was done with pre-sorted categories and so you would have these artificial categories of very concrete nouns, like an electric fan, a teddy bear, a ski mask. So, these are kind of arbitrarily chosen categories of things and I was interested in making an abstract print that would be recognized by these algorithms. So the algorithms are trained on real world photos. But my starting point was, can I make something that's within the bounds of traditional art, so an ink print, but then when it's shown to these algorithms, they will say, I know what that's a ski mask, go. I know what that is. That's an electric fan. So, that was my starting point and that's still what I do. Now, one of the things is that the technology has evolved and so more recently, it's not so much these categories. So, because natural language has moved so much in the last few years, some of my more recent prints look at more ephemeral categories. So, for example, one of my recent prints was raindrops on a windshield. I think we all recognize what raindrops on a windshield looks like if I were to do kind of a sketch of it. But this is another kind of visual category that you might not think about in the same way as teddy bears or electric fans. But it is something that the computer, if you kind of prompt it in the right way, can make an abstract print of in just a few colors and kind of boil down and I guess that's the other side of the coin here is that what I'm interested in is abstraction. A lot of the AI art these days is going into (or a lot of people) we're working with what they call AI is looking at very realistic renderings or making something that is kind of photorealistic. But I've always been interested in the other direction, which is, can we make something that's very, that removes some of the details and that that looks to us like an abstract sketch, but it's still very recognizable by these algorithms.

Well, that's the thing. Your art is minimalist and the other example you're referring to is, I've seen things like by Janelle Shane, in particular, where she uses VQ-GAN and CLIP and the CAN is generative adversarial network and the idea is kind of like running the image recognition network backwards to say, what is the essence of this tag? What is the most cupcake image you can show me? and it's, as you say, very photorealistic. Your art is different. For people who are listening and obviously can't see, we'll have to Google this later, what artist might this most remind them of? Dali, Picasso, anything in particular?

Maybe a Picasso sketch, where it's just a line drawing. I've done a lot of things like client drawing something with just a few colors. A lot of the times or almost all of my work is just a very few colors, maybe because I'm doing things with ink, so I might just do for example, electric fan was just a purple and a black on top of it. I think Calder might be another example. Well, because I use a lot of curved forms in my work, not a lot of angles or straight lines, but I actually hop around in my style. So it's hard to pin down one particular style. Because it different times I might be interested in minimizing different constraints. So, one way to think about your example, with Janelle Shane's work is that I'm also interested in what makes a teddy bear but I'm interested in well, what if you only have two colors? Or what if you only have three pen strokes? or what have you only, if you kind of boil it down to some sort of constraint that is? That makes it a more abstract work.

And it's that abstraction it's quite untethered from the real world, it's a sort of thing that you might imagine a toddler creating, except it hasn't got the sort of things that they stopped doing early on even recognizable legs and arms and such like; it is this this pure abstraction. And yet you've tweaked it so that it triggers the AI to say this is exactly this thing that you aimed it to be and, for instance, you've got Mustard Dream, an art that is classified by is it every engine, you run it through as explicit erotica?

So, there's a few questions to answer, I'll unpack that. One is just the idea that you can make kind of not safe for work, abstract imagery. So, it's not clear kind of, from the beginning that is a concept to algorithms, but it actually is and so after making electric fans and ski-masks and cellos and other things, I actually started thinking about what can we make more abstract categories. So, Mustard Dream was part of a series I did where I was trying to make not safe for work erotic imagery that really looked very abstract that no one kind of recognized. So, in all my work, it's hard to say whether it works on all algorithms, because machine learning is evolving. That one, specifically works on Yahoo's and Amazon's and Google's algorithms, and though it was made now three years ago, continues to work on them and in machine learning, we call this generalization. So, if you have something that works, both on the training set, but also on things that hasn't seen before, which is usually called the test set, then you say that the system generalizes well, and all of my work is intended to generalize well. So, one good example with Mustard Dream is that I did have an exhibition of Mustard Dream a couple of years ago, and it was hanging in the gallery and I took a picture with my cell phone and went to post it to the Tumblr blog and Tumblr said you can't post this image because it you know, I think I forget how their wording is something like it doesn't meet our friendly guidelines or something like that and so that's an example of generalization because I hadn't tested it, in that case, on the Tumblr

algorithm, but it works on the other ones and then sometimes it generalizes to new algorithms as well.

And anyone listening to this should not fear that they're going to find anything that offends their sensibility by looking this up. I was showing it to a class yesterday and I said, I'm not afraid of anyone turning me in for showing you this and I'd done that after an example of how an AI in the UK that the police there use decided that pictures of sand dunes were erotic and I said Well, look, you can kind of see what it might be getting out here, right the shapes and so forth. But then you look at Mustard Dream and there's just no way a human can extract the concept of nudity from that human nude forms and what particularly intrigues me is how you've got this recognized as such by three different proprietary I believe engines and yet usually those things are quite sensitive to input conditions. The idea that they would agree that much about something means that you found something that they have in common, even though it's not what they were trained on. Is there something about that you can speak to as how it does that and how you managed to do that with these engines that I am assuming you didn't use or if you did let us know in the generation of this.

Yeah, that's a great question, a couple of things to unpack versus how do I get this to work, especially on systems I don't have access to. So, I don't have any special access to Google's algorithm other than I can go to their API like anyone else can and I can test it. And I'm going to refer to a branch, a subfield of machine learning, called adversarial images. So, adversarial images is where people study these kinds of trickery that you can do. So, you don't want your self-driving car to see something and think it's a yield sign and really it's a stop sign. So, this is a security problem if you have these machines making decisions for you and so there's a subfield of machine learning called adversarial images. Within adversarial images, what we're getting at is a couple of things. One is that this is a blackbox attack. So, if I have access to the system, and I'm able to sort of figure out its inner workings, and then I give it something, then that's a white box attack, but mine is a black box attack where I don't have access to the system. That's inherently harder and then what's one level harder and I think you refer to this as actually to make it not a virtual image, but a real image. So, as I alluded to, at the beginning, when you have a real physical image, a real physical item, and you take a picture of it, the angle, the lighting, there's all kinds of things that could that could influence what it is the algorithm sees. So, the answer to how do I get that to work is that in the software, it simulates different lighting and angles when it's building it. So, it's not just one image, it's kind of a distribution of images, similarly tested on a variety of algorithms, not just one and arguably, I think this is so and adversarial images in the literature, you can find, for example, a picture of a panda bear that an algorithm thinks are a car, and you can't tell any difference. But it's very easy to trick one neural network, it's very hard to get multiple ones and it's very hard to trick multiple neural networks, including ones in the future, or ones you haven't seen and that's really what my goal is, because I think, then I've arguably really hit on an abstraction within these neural networks, so, there really is something about the image that's evoking some quality of the data that it was trained on, independent of the dataset it was trained on. So, Amazon is, is training, it's not safe for work imagery on something different than Google, it's something different than Yahoo, and none of

them are releasing that. So, yeah, that's what my goal is and so I was very happy that that worked. I want to talk about one second thing that you mentioned, which is how you said, you felt very comfortable showing this to your class, and that there and when I was starting off, I also felt the same way. But in the process of exhibiting this, I've had a number of corner cases that I think are interesting. So, one is even naming it, I named it mustard dream because I wanted to kind of, I didn't want to call it pornographic, or pornographic. But I think I actually did the artwork a disservice by doing that. And so, some of my more recent work, I am more explicit of what I think the work is and one of the reasons for that is I can't really control where this work ends up after I sell it. So, I had a gallery sell this work, so it ended up essentially in a children's hospital and I was wondering, is this appropriate to have one of these images at a children's hospital, they probably don't know what it is and another thing to say is even I have built sort of uneasy travelling, if I'm travelling and doing art shows, I'm travelling with something that essentially the algorithm thinks is pornographic. So, as a thought exercise, you could also imagine that I could at the same time, say, well, not only is this pornographic, but if you run an age classifier, then the age is below a certain age, and then you start getting into child pornography, where they're very strict laws and I not bringing this up, too. I'm not bringing this up, potentially to be provocative. But I think even if you say, well, there's nothing I recognize the fact that you're travelling with something that in certain parts of the world, people aren't going to understand and it might trigger a system that you have no overt control over. It's not always as black and white as you might hope it would be.

Well, I can see you travel quite a lot. You've had exhibitions in New York, Shanghai, Bulgaria, Zurich, New Delhi, Montreal, and it has probably occurred to you the possibility of going through a border check, and they scan your computer with an AI to see if you've got anything on there, they are interested in?

That's true or it could be that the artwork itself is displayed and there might be some automated system that says that. So, it's very hard to argue, I think people know, against these algorithms and so I will, I wouldn't want to find myself on the other side of that argument.

And that really exposes something else that I want to get into and I'll point out by the way, that your example of the Stop sign being mistaken for a Yield one is not random, there was a paper that did exactly that by attaching some carefully placed tape to a stop sign that so that it was recognized by the algorithm as a yield sign so. But now, that takes us into the category of cybersecurity, and I had not previously thought of art and cybersecurity overlapping this much, but that's clearly where you're going, or have gone and is that of interest to you? Is it describing how that intersection is affecting you?

I mean, it's certainly something that has come up in my work, but I don't have cybersecurity necessarily front and center in my sights. So, I have done work, for example, that triggers a facial recognition algorithm. So, I've made abstract prints that trigger face detectors. I even did one that triggered specific identities. So, I did a collaboration of an oil painting and the oil painting is abstract. But if you feed it to an algorithm, it'll think it's Nicolas Cage. So, it's sort of doing a very specific facial recognition, using kind of the latest facial recognition algorithms, so I think it

is possible to bump up against these kinds of cybersecurity concerns, but really, my interest is more is more what is it that these machine like? What is it that these machines see that make are causing them to make these conclusions? Like how is it that they're seeing the world and if you make an artefact, or you make an object that they understand, then that's a way of kind of getting in their head and understanding, how it is that they break down imagery.

There seem to be the plots of several novels and movies of the week potential here. Where artwork ends up with Nicolas Cage being arrested for being in the wrong place. And as you say, you're labelling yourself as an artist, you're not going into that cybersecurity realm. But clearly, cybersecurity people studying this interested and it's exposing important things that this new thing that we, as you say, you call it alien, but we're trusting it, it has this view of the world that we don't understand, and you're showing us how much we don't understand it.

That's right and oftentimes, the ways we don't understand it highlights very specific issues in machine learning. So, one, I'll reference a couple of prints and then I'll talk about how this worked. So, in one of my prints, it's called Cello and it's an abstract print with red ink and black and there's these you can see a cello shape in the foreground and you can kind of see the shape hovering behind it, which is the cello player and already that's one issue in the picture, and they don't differentiate this this so, if you train an image, I call this the dumbbell phenomenon because it was pointed out by the Deep Dream researchers that if you train an image on dumbbells, you often will get arms attached to the dumbbells because it was just going to it thinks there always go together.

It's never seen a dumbbell without an arm on the end.

Very rarely sees a cello without a cellist and so if you ask it to draw a cello it's going to tell us a little bit more subtle is that if you look really closely at the at the cellist and you look at the training set, you'll see that the cellist in the image that got drawn is left handed and almost all the cellos in the training set are right handed. And the reason for that is a very peculiar thing in, which is the machine learning systems are mostly blind to left-right symmetries. So that's not something that people think about that much. But they are trained, because it makes them a little bit stronger to train on reverse images that are also reversed. That's a standard part of training, and so you do sometimes if you look closely, it's kind of a nerdy thing to point out. But it's something that I noticed and it's something that kind of bubbled through in the art organically on its own. There's a second one, which was in the same series called Measuring Cup and when the systems draw the shapes, they also choose the colors and for measuring cup, it drew what looks like a measuring cup, you might use it the kitchen. But then he chose this really bright green for the color, which was really odd, because all my measuring cups are clear and have red or blue late lettering on them and so I looked at the training set to try to understand what it is this bright green color was and it turns out, there's something called uranium glass, or that was a depression era thing, which is collectible now, so at some point, in the '30s, or '40s, I think they were putting little bits of uranium in the glass, because then it would glow green and then these became collectible cups and then people sell these things on eBay and so long story short, when you go doing an internet search for measuring cups, in the normal set of measuring cups, you also get a bunch of these bright green measuring cups, and they're over-represented. I'm

describing the situation in pains because this is a something that's a common issue in machine learning called sampling bias. So, you have something that's rare in the real world, but then when you collect data, it's actually quite common. And so this idea of what a measuring cup is, has been kind of polluted by the fact that these collectors are more eager to post their images online and so they end up filling up the data set and then this algorithm actually thinks that measuring cups more often than not, are bright green. So, a couple of really specific hyper specific examples, but these are things that are part of machine learning and they're kind of obscure, but they kind of bubbled through and they influence how the algorithm sees things and how the art gets generated.

Have you ever had a case when you were going for abstract and it came out and it would be recognized as what you wanted it to be recognized as, but instead of looking abstract, it actually looked pretty much like something different. That ever happen?

Yes, I think that's happened some with some of my recent work, where I mentioned the raindrops on a windshield, I did a series of these and I think for the most part when that happens, it's not as interesting to me. So that happens sometimes, and maybe those will end up in my art. But to me what is always more interesting is when something comes out that defies my expectations in maybe a way that's not obvious. So, another one in this series was called shatterproof glass and so if you ask, what shatterproof glass looks like you would probably think about your car and you would think well that's just a pane of glass. But if you do a Google search for shatterproof glass, you will find it's almost always shattered. Like someone has hit it with a hammer and so, the idea that or it's not shattered but it's you know it's in pieces or it's been attacked, because that's when you go to buy shatterproof glass, that's how it's advertised and so I thought it was really interesting that the way we experienced shatterproof glass is very much not what a machine learning system would think of as shatterproof glass and that was one of my fears. So, I think the answer your question is that does happen sometimes. Sometimes it ends up being what I do in the artwork; but more often than not, I'm interested in these kinds of surprises that defy my expectations.

You've run a number of exhibitions where your work is exhibited on walls for people to look at. And in a real, or let's say hypothetical, situation. You're at one of these exhibitions, you're standing behind people who are looking at one of your works and what can you imagine them saying that for you really exemplifies the kind of reaction that you would like to get.

So yeah, it depends a lot on what I've had different types of exhibitions, some of my exhibitions, I'll have completely themed around a single object. So, for example, I've done an exhibition of just knots. Knots is kind of a funny one, because it's not so much a particular color or shape, but it's more the relationship between shapes that I think make something a knot. Certain curves and the way that these curves intersect with other curves and so I think that, what I'm what I hope people do when they see this work as they might actually think of the objects themselves in a new way, the computer has a way that it recognizes the knot and by having the computer draw 5, 10, 20 knots, you can actually start to get an understanding of its internal logic of what to computer makes it a knot. It's light, and it's dark, and it's relationships. In other cases, we talked about the not Safe for Work imagery. That one is similarly I think it's interesting, just where it

might be a little bit more abstract. But it's interesting that some that the computer has a concept of what is erotic imagery, or adult nudity, or whatever these systems are trained to filter on and it's also very interesting that it's so foreign for us. So, I think in those situations, in different exhibitions, depending on what the subject is, I think people would go away with different things and sometimes when it's concrete nouns, knots, or ants are these things that are you run into every day, you might think of those objects differently the next time you see them in real life and in other situations where it's something more a spam filter, you might have more respect or more wonder or more confusion about how these things work, because you can actually see some of the things when they're asked to express what they're looking for, you get kind of surprising results.

And you're touching on something there that I want to ask about and this might be either how do people view your work at these shows? Or how are the shows promoted for people going. Is it for more gaining an insight into artificial intelligence or gaining new insight into our world?

It might be useful to start off that by saying how they're *not* promoted because it might be a nice foil to set it up that way. So, I think what people want, or what there's a hunger for, or what there's even a market for, is this idea that there's a computer that is an artist, so you have this sort of autonomous algorithm that's off creating things on its own, and isn't it wonderful that we've kind of taken creativity and put it in a box and it can go off running, that is not how I try to market this, then not how I try to present it, what I try to do, as I tried to say, look, I have this new perception capability, this alien way of seeing things and in fact, it's a way that we built we've engineered it and the artwork is kind of like you would expect if you were exhibiting art from a certain culture, from a different part of the world or different sort of subculture, then that's ideally how it's presented that this is work done by a computer and recognized by a computer, but kind of with me in the driver's seat. So, one way to think about it is this kind of a collaboration where I'm picking the subjects and I'm picking kind of the themes, but it's really the computer. It's the neural networks at the end of the day that are kind of deciding how to configure what ends up going on the page so that it's something that they recognize.

And you've exhibited all around the world, any breakdown by geographical region, or culture in terms of the reaction or is it a universal reaction?

I think it's fairly universal. My work is very abstract visually, so it's abstract representational. I think that there's certain parts of the world that are more technically savvy, and I think there's some certain personalities are more open to the idea of machine perception. But I think we're all kind of curious how these systems work and I think that, facial recognition and the idea of self-driving cars and the ideas, these exist, systems exist, this permeated our culture enough where people are used to even on my phone, and my phone is a mess, because I'm taking pictures of my artwork all the time and then when I go, I don't know if you ever looked at your, my iPhone tries to classify for me here's the food you ate, here's the animals you've seen and of course, if I go into my animals, it's artwork of a killer whale and it's kind of confused as to what I'm into now and I think that the way that these things intersect with each other, the way that these recognition systems intersect with the recommendation engines, and our kind of day to day life is part of the surprises for me.

Fascinating, where do you plan on taking this next?

I'm continuing so lately, the big changes over the last year is, that natural language has kind of met up with computer vision. So, it used to be the computer vision was very categorical and if you wanted to recognize a cat from a dog, you had to build a cat versus dog data set. In the last year, I've been looking a lot at more natural language categories the raindrops on a windshield and so I am tracking a lot of the developments in machine learning, a lot of times track the updates and machine learning, and there's not a change. So, one big change that happened these past year is that internally, neural networks used to use something called a convolutional neural network. But they changed [to] using a different technology called vision transformers. So, vision transformers are based on natural language. It's very nerdy but it's been a big change within the machine learning community. And so I updated kind of my software expecting that these vision transformers would see things very differently than the systems that came before. But actually, they don't and by and large, they make they draw things that the systems before them still recognize. So, sometimes, I don't know what to expect myself and it's just kind of adapting to the machine learning trends, and then seeing how the artwork evolves. Which is kind of my goal is that over time, I'm kind of building a body of work that reflects kind of this path that we're all on with these machines are getting more capable and more advanced all the time and embedding themselves in our lives in different ways.

You were talking about the progress from just categorization, and there was something released in the last year where they had an AI where you can give it a description of what you want, you can say, show me a pig drawing a pig riding a unicorn, and it will do your picture of that. That's pretty big step or at least it looks like it.

That's a huge step. So, the big the big step there was about 14 months ago, OpenAI released two technologies, they released the first version of Dali, and they released a system called clip and it was really clip that was, in my mind the surprise, it was a system that instead of being trained to base on a data set, it was based on just kind of a big corpus of internet text and it tried to match images, score images versus the text that they were situated with. So, you could think about it as kind of a captioning system on steroids. So, when you see an image and a caption, you can think well does that image match the caption and they built a system which would kind of play this adversarial game or play a game with itself and tried, to given five images and five captions, which one matched the best and it got better and better at that. So, yes, as a result of getting really good at that task, you can feed in a general caption or phrase and it can tell you given a set of images which of the images score best against that and then you can use that as a kind of steering wheel or you can a steering wheel where you say okay, well if that one's close, how do I get closer and you essentially turn this into a drawing system.

Just amazing. To get near a conclusion here, what would you like to do with in the next few years, what big steps would you like to make in this art form? Assuming that's where you want to keep going?

Oh, definitely, I think I'd like to keep building is more or less in the same way. So, I'm going to continue making friends. I have done some digital work. So, NFTs had been a big thing this past

year and to me that the interest there's a market for doing digital work. So, I have experimented with kind of bundling bits of the neural network's logic in the graphics and so that might be a direction I push out a little bit more. But I think the bread and butter of what I want to keep doing is kind of pushing the boundaries on these systems as they evolve. So, a good example would be most of the system's, all the systems, I think that I've talked about up to now, really, they don't experience the world in any meaningful way. In fact, they never even see any motion. So, if you and this is also true for CLIP, so the very latest system, so if you have a system that knows what raindrops on a windshield looks like or knows what a cat looks like, it's never seen the cat move because it doesn't view sequential frames. So, I have done some experiments, where I do hybrid, where it actually puts these things does like a little stop motion animation. Because I think that looking at what constitutes motion is interesting for computers and vision systems. But I think that just kind of speaks more generally to as these systems evolve and as they learn to recognize different things on our behalf. I'm interested in attacking, and that might be getting hyper specific. So, I think it'd be interesting, for example, to get in the proverbial head of a self-driving car, and how does it see the world versus your home security system. But for now, I'm using mostly research systems that are being released and I'm feeding the drawing system these systems, and as they evolved in the imagery kind of evolves under it.

Well, and that that is really interesting. You were talking about motion there. I remember, how it took quite some work to convince self-driving cars of the concept of object permanence. That things were still there when they moved behind something that obscured them from view for a while. This is great. How should people follow what you're doing; any exhibitions coming up or sites you want them to look for? Or anything else that's your output that we can refer them to?

Well, thanks. Yeah, that the easiest way to follow what I do is probably just to follow me on Twitter or Instagram. My handle on both is dribnet and my work is all posted up there. On Instagram, I guess if you just want to see the work and on Twitter, if you want to hear my snarky remarks about it as well and then if you want to just go to my website, that's drib.net. There I kind of listed some of my recent exhibitions and press and normal stuff like that. Yeah, I continue to those will be where I continue to push out what it is I'm doing.

Fantastic. Tom White, thanks for coming on AI and You.

Thanks for having me.

That's the end of the interview. Isn't it amazing how AI is causing us to think in new ways even about art? Plus all kinds of implications for steganography – hiding one image inside another. As long as AI doesn't see the world the same way we do, it can be fooled, and people could take advantage of that, by, say, making an innocent image look pornographic to AI and getting someone in trouble for it. Or, of course, the opposite. It made me wonder what it will take to get AI really seeing the way that we do. How much are we going to have to understand about the human visual cortex to do that? Go look at Tom's works – there are links in the show notes and transcript – and think about what these mean for how AI is seeing the world.

In today's news ripped from the headlines about AI, and since we're on the topic of art, there are now two state of the art tools doing unbelievable image generation from text prompts. First there was DALL-

E 2, from OpenAI, and now there is Imagen, from Google. You just type in what you want, and they generate a photorealistic picture of it, no matter what it is. You can go to openai.com/dall-e-2 and see its images of “a painting of a fox sitting in a field at sunrise in the style of Claude Monet,” or “a bowl of soup that looks like a monster knitted out of wool,” and all kinds of other inventive mash-ups, and they look as good as you could imagine anything being that was designed to fit those prompts, and yet of course DALL-E 2 is not finding images that have already been created, it is making it up. And now Google has Imagen, which does similar things, only with different technology behind it, and they say that it is rated better than DALL-E 2. In fact it’s so good that they are not letting just anyone create images with it, because they’re concerned that people will create inappropriate or illegal images with it. Use your imagination. But I would expect that eventually it will become available anyway, just like everything does. It’ll put a lot of Photoshoppers out of business when you can just describe the image you want and get it back in seconds.

Next week will be our two-year anniversary, and so I’m going to be doing a special episode just explaining some of the terminology that gets thrown around on the show rather casually now. Words like AGI, and Alignment Problem, to help those of you who are a bit new to this so you’re not feeling left out of the club. 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>