AI and You

Transcript

Guest: Sathish Sankarpandi

Episode 63

First Aired: Monday, August 30, 2021

Welcome to episode 63. We are all about making connections on this show. We look at AI from so many angles that you might think of the show as a recreation of the way that neural nets are trained to learn something new, by showing it a zillion examples of a cat and then it learns what a cat is. Maybe that's how we're learning what AI is. Although AI is far more brittle at learning by example than we are, which is why it's only currently good for learning what things like cats are and not what things like AI are. But it's getting better all the time.

Today's guest is Sathish Sankarpandi. He is head of data science for Orbital Global, in Suffolk, England, and he is the mastermind behind Orbital's VirtTuri digital avatar. That name comes from a contraction of "Virtual" and "Turing." So we're going to be talking about AI customer service agents. But that's not all Sathish has done; let's find out more as we get into the interview.

Sathish, welcome to the show.

It was really great to meet you and speak to you. It's a really great opportunity. I'm very excited to be with you.

Thanks, So, you are in a position at Orbital Global in the United Kingdom. Is that correct? What are you doing for them?

So, I work as a head of data science for Orbital Global and a company called VirtTuri. Basically, what we do is we have lots of AI products that's going into predominantly healthcare domain from digital assistants, then we have a little bit of dashboard projects which would help, plan and deliver any sort of health care campaign. And currently the project we are working on is funded by Melinda Gates Foundation, and we are working with Rwandan government to actually improve their health microplanning and delivery of services. For example, currently we are doing scored vaccinations, so how you can increase their coverage and reduce health inequity in Rwanda using AI.

And is that using VirtTuri or is it another mechanism?

Yeah, VirtTuri is a digital assistant, so it can help you answering questions, just by typing into a free form textbox. But, the whole application is on optimization and planning. So it's actually bringing various different data sources and trying to predict for example which regions you have to target first or where is the, because the disease outbreak is high. Or for an example, you have this certain types of vaccine where you require cold storage, so what are the optimal questions to place this cold storage, so you can actually maximize the coverage and also minimize the amount of cost and time spent. So that's where it is. So that's the whole dashboard, and it is for various

different stakeholders, for ministry of health, for the director level, to the program managers who actually plan them. And to other managers who are doing logistics and others.

So, I want to back up to a higher level here in looking at this because my main question when I hear this is about how you can apply AI in real world medical situations reliably. So that if you were in the United States for instance, you wouldn't be facing possible ruinous malpractice suits for making a mistake; and get an idea of the context there because this is fairly new. And you've been talking about VirtTuri, which I believe is coined as a portmanteau of Virtual and Turing. And it's, as you described, an avatar. And I've been kind of thinking of it as Dr. Siri, or Nurse Alexa, if that's not getting too gender-specific. It's an avatar; it's something you see on a screen. Do you talk with it verbally and does it answer back that way?

Yes. So, VirtTuri has a virtual element, which is an avatar. And the idea is to create a hyper-realistic avatar. So that you actually,feel like you're talking to a real human. So, basically go in there, you just speak to it, you can also type if you want, just speak or type to it, and lots of things happen, and then you get an answer back. So, one of the differences between Dr. Google or some other platform is, you know, this is very niche: it doesn't work on everything, it just works on a particular topic. And it's not a diagnostic tool and it only gives you advice or information on things. For example, COVID, if I want to say, you take a COVID vaccine, you know, you can ask questions about, my arm is sore or, you know, what do I do, and all those things. And one of the advantages of VirtTuri is it's highly engaging and we know from our research that if you see by seeing a video or interacting with someone online, your engagement or your information recall is very very high, or the knowledge transfer is so high. So that's the reason why we have VirtTuri there. And also, it can help with the loneliness epidemic, loneliness that is widespread with elderly adults, across the globe. So, this can play a vital role it can just even saying to the elders that every morning, have you had your tea, it makes a huge difference. So, in a way, you know, it's highly engaging and it also can transfer information accurately and quickly. So that it stays with someone quite a long. And also you can have natural conversation, it's not like you have a pre-written text and it has to match, it's not keyword matching or anything. So, it's very natural, like you speak to another human being.

So, if I go back decades in the history of AI being used in medicine, there was a system called MYCIN, that John McCarthy took apart - in a rhetorical way - because he said this is not ready for use, if for instance a patient has a cholera infection, it will recommend a course of treatment that will result in the patient dying. And that was one of the things that precipitated another AI winter. Now, here we are, fast forward, we're in this situation where you're using VirtTuri and things have changed. What were the challenges that you faced in engineering this and getting it accepted to be used?

Yeah. So, I want to start with saying it's not diagnostic, it's only information providing tools. So that takes out a huge burden or a challenge out of it because, you know, if it's going to be diagnostic, then you're actually kind of replacing a GP or a health professional in that sense. And it is very very difficult to do. And I would say that's why most of the systems fail because they try to do diagnostic element and there's so much of compounding variables that actually create,

you know, a treatment for malaria is different for different person because of age or BMI and everything. But, with VirtTuri, we are not focusing on diagnostic; we're only focusing on providing information or signposting to different elements. The huge challenge here, is actually understanding because, you know, it's a free-form natural conversation you can have with VirtTuri. So, actually understanding those nuances, and how do you make computer to understand that. People can say ask the same thing in various different ways and you cannot train the AI system with various different ways. It's impossible because, you know, I can be someone now and maybe after later, and then I can ask the same question in a very different way. So how do you actually make it understand those differences or variations of the questions? So that was a very big challenge. And us as a kind of ML engineering, machine learning engineering, then actually putting that into production, and it has to be really fast. When someone asks a question the system shouldn't wait for five minutes before it gives an answer, and that's not engaging at all. So how do you make it quicker? So, the response time has to be fast, it has to be accurate and also we have a visual element which is an avatar. So, all of these things has to come together. And there are various different microservices running and all of those things have to play in harmony, so that you get those three things that I just mentioned. So, those are the biggest challenge,s I would say.

So, we're talking about natural language processing, and how do you train the model? What is the core engine here that's learning how to interpret the human words? And how do you train it?

So, as a very simple understanding of how it works is, basically teaching a computer to understand language has been long studied. And one of the biggest breakthroughs was actually what is called as vectorial representation of words. So for each word how do you add numbers? So, for a word, say COVID, there are some numbers observed with it, but those numbers are not just any number; you can't pick any number, because these numbers have an association with other words but in a sentence, it makes a very big difference. So, for an example, "I lost my cell phone," is different from "I lost,my book in a cell," which is a prison cell, even though the cell was there in two sentences, but it means a very different thing. So there's a context around the word. So, there was a biggest breakthrough was, you know, actually trying to understand the context around each of those words. And actually giving those numbers, it's called a vectorial representation or an embedding. That's where it all started. People are trying to, you know, make this a vectorial representation as good as possible, so that you can actually, it can actually understand kind of semantics of the sentence. So that's the rudimentary aspect of natural language processing, I would say.

So, I'm then taking that, so once you have the numbers, then the computers can understand numbers, you know, and then you can feed it to any machine learning system you want, you want machine learning, deep learning. So, for VirtTuri we have a hybrid or a hierarchical architecture which produces a low level, straightforward, simplistic, regression and methodologies. And also we have very complex deep learning methodologies at the end. It works in a way that making sure that the information is accurate, the information given to the user is accurate. Also we have the balance of speed. So that's how it works.

So, there's understanding what the person is saying and then you've got, well, there's telling what they said. We turn their sounds into text; but now we have to interpret that, we have to just put some meaning on that. And then you have to turn that into a question, or then you have to infer from that what to say to them. Are they asking questions or are they making requests, or what is the person saying?

Yes, definitely, you know, that's exactly what the machine learning algorithm is going to learn. So, it's actually, the first step is converting the system numbers and then actually teaching them what numbers are what, you know, is it a request or whether this is a question. And if it's just a question, you know, what question, what kind of question is that, you know, do you have to do it. Another thing is, you know, do you have an answer for it first of all there might not be answers for it, how do you answer that. So that's also a bigger problem. So, taking that, understanding that, and actually figuring out what it is. So that's a big challenge area.

Right. Now you referred to how it's only operating in certain narrow domains, like it might be COVID; but there are other ones. Within any one of those narrow domains, it's much easier to answer if you know that you're only talking about COVID; there's a limited number of things that the person could be saying, and a limited number of responses that you can make. But do you have to first determine what domain you're in, or is that predetermined when the conversation starts?

It's all automatic to be honest, you know. It's so say for example you have various different machine learning algorithms sitting behind it. If you're asking a question about COVID, it automatically recognizes that it's based on COVID and it can give you answers on COVID. But if it's asking anything something about planet Mars or Uranus or anything like that, it doesn't have an answer. So, we have to figure that out as well, you know. Actually, you know how we detect it because the machine learning is only or trained on certain kinds of data. And this is out of scope or out of the domain question, so how do you detect that? So that's also another other system that is working continuously trying to understand what exactly they are saying.

Right. Now you're saying you went to construct a hyper-realistic avatar with this, there's a face to VirtTuri. Is it more than one face or is it just the same face?

Yeah. We can have more than one face. Yes, definitely we can have more than one face. At the moment we just have one, that's only for start, but we can have any number of faces we require.

Can you compare that the way that's done to some of the other avatars that people might be familiar with like Samsung's Neo or IPSOFT's Amelia; do you know if you're doing that the same way or differently? Or how is it done?

So, the biggest difference I would say is, you know, ours is an audio-driven avatar, and I don't think many software companies are doing it, it's completely audio driven. So, usually what people do is they do deep fake, so which requires a video, so you have to have a video or a reference video, and also you have a video which you want to translate into. And then that's how the deep fakes work. But what we are doing is, there's no face at all, and then it's all driven by

audio. So, even though they type in a text, you can convert that into audio, or vice versa. And then it's all driven. And what happens is there's a speed synthesizer which is trying to understand what the speech is or what the audio signals is. And then this is given into a deep learning model, which will convert all those things and it will just map out different actions for different points on your face. And then we render that. And this has to work really fast.

Is the building block of the facial movements something called a viseme?

Yeah that's right, absolutely, yeah. So, for a face, there are various different musculoskeletal joints whenever you're speaking something or you're showing some emotion, there are certain things that happen to your muscles, and your bones. So, these points have been already mapped by various different researchers, and you can find them all over the place, all over the world. And taking that, and then actually converting or finding audio signals that can control those points is what this whole subject is about. And visemes are one of those things. Visemes from the audio signals is mapped to the face. And then you render that, so that you get a hyper-realistic motion.

So, is this in the same class as deep fakes? There's a famous video, now from a few years ago of Barack Obama, and it's driven by the audio, in this case, of Jordan Peele saying some things in this facsimile of Barack Obama's voice, and then the video of Barack Obama looks like it's saying these things, and it was driven by what his voice was. Is that the basic technology that you're using or have I got it wrong?

Yeah. I think a deep fake is completely different because deep fakes, you require someone not just speaking but also a video of that person. Like say for example, I can speak, but my video has to be recorded. And I can convert that but I can overlay anyone's voice to it. It's very different. What we are doing is, you don't require any video at all; you just have to send the audio signals in. That's it. So, the input is audio signal, output is avatar. But whereas in the deep fake, you should have video input, video and audio input to it, and then you get a video output or an avatar output.

So, does the avatar's image synthesis, does that have an underlying model of musculoskeletal physiognomy and kinematics?

Yes, of course. It's all been studied very well. And what you have to do is you have to connect the linguistic element and then the muscle, so which for what sound, what muscle activates. So, that's the whole study here. It's not straightforward.

Yeah. No. That's a huge amount of work to go to; and so for someone who's thinking, was that necessary, what value you add to the conversation from creating that hyper-realistic avatar, when the conversation could have been a disembodied voice like Siri or Google? Why go to that much effort to add the face?

Yeah. So, actually there's a big research which is done over with 10,000 people or so. So, they compare text versus audio versus image or video to see, you know, how much information is retained or how engaging is, how long the users where with the screen. And what they found was, they found about 70 to 80 percent more people can recall information. And it seems they

know they can get engaged. This is very similar to one of the works that we did with MySpira, which is an asthma application that we created from Orbital; it uses something called augmented reality. So, it's a gamified way of teaching people how to use, ask my inhaler correctly. And what we found was up to two to four times, people can actually understand the process of inhaling or using asthma inhaler properly. So, I think it's about the information retention and also engagement.

Right.

And if you take older adults, this becomes very necessary because it creates a different feeling or different sensations so that because they think that they are actually speaking to a real human being.

So, you've jumped the uncanny valley?

Kind of, yeah, that's where we are. I mean there are little giveaways, you know, for certain sounds there are certain giveaways that you can say that's an avatar. But first for others it's very good. We were even thinking to have a little label that to say that you're actually speaking with an avatar because people shouldn't get the wrong idea.

Yeah. Same here the Google Business Assistant or whatever they call it, automated telephone system has to tell you that it's an automated system when it calls you in, it called me a couple of times but it's very good. Do you customize the look and sound of the avatar according to who it's talking to?

Yeah. At the moment the functionality is not available yet but, you know, the idea is like it's transferable, all these things are transferable. So, you can transfer it to any ethnicity you want, even we can let users to pick their hairstyle of avatar, they like, or who they want to speak to; male, female, anything they want to do. And also they can also change the voice of avatar that's speaking out.

And is there any affective computing in this? Does it either recognize emotion of who it's talking to or generate any depending on context?

At the moment it doesn't do it. It can detect emotions from the voice, basically. But what we can do is we can incorporate the camera, of their device, of the user's device. And then we can also detect those emotions. But the problem is privacy; privacy is the biggest concern so… And that's the reason why we didn't go into camera, but camera can provide lots of useful insights as to, you know, what the user emotion is than the voice, of course. Yeah, but there's an element of privacy there and that's why, you know, we said we are not going there. But it's quite possible to do it.

This is not the only AI application you've done in medicine: you were describing something before we started the interview about a project that would measure the likelihood of people falling accidentally. Can you describe that?

So, in the UK - I'm just thinking about UK, but it's a worldwide problem. One in three adults who are over 60 years fall at least once a year. And that can be really fatal at sometimes. People have lost their hip, hip fractures, hip injuries and the other thing is psychological effect. If they once fall, they always have a fear of falling again. So, which means they are not mobile at all. And that actually impedes their independence. And that is going to have a worse effect on them. And then they're going to fall, anyways fall. And it's costing a lot of money about 2 billion spent, just in the UK…

So this is an inportant problem.

It's a huge problem and it's a worldwide problem and it's been well recognized and that's for over 60 years, and over 80 years, one in every two adults fall. So, it's a huge problem. There are two big elements of false risk that can happen. So, one is an environmental hazard, so that's to do with trips, any hazards they have in their home, like loose wire, a loose carpet, bad lighting and all those things, those are environmental effects. There are lots of interventions, or there are lots of things that are being developed by government and other agencies, such as Age UK and other Nice UK and all those people. They have given a little questionnaire to check their house is clean and they are free from environmental hazards. But the biggest problem, I think is the intrinsic effect or what is happening within them. So, loss of balance because of, for muscle strength, the other thing is, you know, their gait, there may be an issue with their gait. All these things are intrinsic and these are all much nuanced to find at a very early stage. So that's very difficult to detect. Currently what they do is they have the health professionals have a questionnaire, which they actually go through or they ask people to walk and they see it in there with their naked eye, and they just score it. But it's very subjective and depends on what time of the day you're doing and the laboratory effect, white coat effect and all those things come into play.

So, you basically did the scoring with AI.

Yeah. You ask the users to wear the sensors, and ask them to do the same protocol that they're doing, like walking, sitting down, and all those things. This algorithm can detect all those nuances that the healthcare professionals can't see that they're with their right. I mean even at a very early stage, we can just detect that someone is going to fall soon or they are in the slope of determinant, you know, or their health is declining in some way. So, it is better to make an intervention, it's just a screening tool, it's not, it does, it's not saying that you have this problem, but it's just a screen screening tool. But can give you a very in-depth or, you know, it can give you different variables, like where the problem might be, whether it's on the lower body or upper body, you can also detect, you know, how use your heels right and everything, that could be a problem for someone, maybe they have to change your shoe and everything. So, it can give you all those things, you know. When I did my PhD, I kind of,got about 90 different variables that you can actually see where the changes are happening. So, you can see how nuanced.

It's doing this by observing data about how the person is moving and interpreting that…

Yes.

And was it a supervised learning model? Did he train it up on human assessments or how did it learn?

Yeah. So, it's a supervised model, of course it's a supervised model. So, you can only say, so what we did was we had fallers and non-fallers, and then we just gave everything to a deep learning mode system, that's one way of doing it. So, give it to a deep learning system. And it's a black box. Okay? So, you get whether they fell, whether they're going to fall in the future or not or how high is their risk, you can get that. Another way to do is use a traditional regression analysis or anything but there, what you can do is, it's much more interpretable because you know what sort of variables you're getting it out, like say what is your walking speed, what is the range of movement of the rounds, you know, you can exactly get. And you can just simply say whether it is in the normal range or not because you can get the same data from a normal people, for healthy or, you know, who are of the same age but you think they are healthy because we have done all those further assessments or detail assessment via a gold standard method. So there's a reference and you can just simply say that okay, they're going to fall because I can see lots of red in these variables, because it's out of range or out of normal range. So that's an idea there. So, there are a couple of ways of doing it.

It seems to me there are many problems that you could apply that same model to. For instance, cognitive impairment is something which at least here in Canada, the tests for that are expensive because they have to be human administered cognitive tests repeated over a period of time. And I don't know why that couldn't be done by an AI. I think that would be easy. The same kind of training process, you're talking about falling in one case and you're talking about brain cognitive failure in another. Do you see that kind of application taking off for what you've demonstrated here?

Absolutely, I think we have one project that is about to start with an Orbital itself. It's for actually measuring brain imaging or brain problems, like something to do with your neurological problems with their brain, so that, you know. And how you can do that is instead of using sensors that are placed all over your body, you can just take an EEG measurement from the brain. And then, you know, you can just do the same thing as whether they are normal or abnormal. And then you can intervene. You can also extract very many different measures as to, you know, say. And you can personalize or tailor the treatments based on what sort of results you're getting out of it. So, which variables areout of range or, you know, it's below or above the normal range. So, you can make interventions or tailored interventions on that. And this can also measure the prognosis, like say you're giving an intervention, you can also measure them. I mean ask them to come back and see, you know, whether that intervention is actually having an effect on that person. If not you can change it.

And thank you we have a neuroscientist coming on the show in a couple of episodes, Olav Krigolson, and he looks at EEGs all day. So, I'll take it up with him. Just to jump here in the remaining time, you also have done something with a fascinating title of the Ear Worm Project. And ear worms, I think of it being those annoying jingles, like the… I don't know, *Banana Splits*

theme or something that you can't get out of your head, once you've heard it. What was this about?

It was an interesting problem and they want to study and their neural neurological ability or neurological levels with people and what is the connection to earworms? So, the idea is like if you can repeat the earworms in the same tempo as the original song, then you have a neurological system in your brain than if you can't, you know. It might be a problem, it might not be a problem for someone but, you know, if you can't repeat it in the same tempo, it's good, if it not, you know, there will be lots of there might be a detailed assessment or further assessment is required for the person. Or it might give away to… there are some health issues with it, you know, let's call it that. So how do you actually find out whether someone is hearing is on the same tempo. So that's a simple question. So, the way we do is, so I was using accelerometers to actually find out elderly people falling. So, we use the same accelerometer and ask them to wear it on their hand and tap it for the tempo. So, what are they hearing whenever they hear the earworms? We ask them to tap it. So that is the tempo that we have to get. So, I was a person who actually wrote those algorithms to figure out those tempos from the tapping because some people can tap it slow, or really hard, or it can decline, it can go up, that's a big signal processing element there. And we kind of figured out how to actually get that and standardize all the whole algorithms, how to figure out this tempo, and was very helpful. And if you saw from the paper and it was covered.

So, this experiment is playing subjects, I don't know, commercial jingles, and they're cursing, but you're then measuring how well they can reproduce what they have heard, how much it's stuck in their head; and what's the application of that?

So, it has lots of health applications towards it, you know, I was only involved in the algorithm side of it because I was not involved in the health aspect of it, it was a research that is done by Goldsmiths University, and they didn't know how to do the tempo thing. And we came up with this idea that we can use an accelerometer and we can do it. I am not aware of this thing used anywhere but I know it has lots of health benefits if you can reproduce the tempo of the song that you heard earlier or the jingles you heard. And it relates to memory impairment, cognitive reasoning and stuff like that. So, it's connected to all those things. How it can be applied ,and all those things are still unknown.

Is it more purely aspect at the moment in neuroscience but fascinating stuff. I mean it's just one of those things that who would have come up with that. And so that's really interesting to hear about. Now just drawing to a close here, where are your efforts going next? Is it further development of VirtTuri? Do you have another project? What's your next challenge?

So, yeah, we are definitely further developing VirtTuri, as to increase or improve its emotional intelligence, so that's where the whole thing is going. And also there are certain giveaways in the hyper-realistic avatar generation to certain words and things like that. So, that's where we are moving with VirtTuri. But there is huge number of projects as I mentioned one was actually figuring out their brain activity or how bad their brain was injured or anything after they've been in injury. So that's one project. The other thing is we are working on heavily on optimization and

planning for healthcare deliveries in underdeveloped countries or low to medium income countries. So that's a very big project that you're working with Melinda Foundation, Ministry of Health, in various different countries. So, the plan currently is we have got this tool working for Minister of Health for Rwanda. But the next thing is scale up to 74 countries who are LMIC's, the low to medium income countries. These are some of the projects that we currently working on. There are probably so many other things, there's projects on pharmacovigilance to identifying adverse events from the text, to name few, there's quite a lot of things going on.

And these are all fascinating applications of AI that are demonstrating just how it's changing things on an accelerating scale. How would anyone listening who wants to find out more about the things you've just been talking about, how can they learn more about those or what you're doing?

So, they can go to their website, just http://www.orbitalglobal.com or you can just contact me from my email address with the sathish@orbitalmedia.com.

It's been fantastic, a real pleasure having you on the show, Sathish. And good luck with what you do with VirtTuri, and let's hope that gets picked up in even more markets and that work expanded.

That's the end of the interview, and I think it was very good for getting a finger on the pulse of where we are with conversational avatars at the moment and how they're getting out there. Orbital doesn't have as much money as Google and yet here they are doing something with avatars that only a few years ago would have required Google-class resources to tackle.

To update you on one of our guests, Professor Ryan Abbott and Dr. Stephen Thaler have been granted a patent for a novel food container design on behalf of its AI inventor, DABUS, the Device for Autonomous Bootstrapping of Unified Sentience. South Africa granted DABUS the patent despite prior denial by the US, UK, and European patent offices. You heard from Ryan when he was my guest two months ago, but the patent was only just awarded; the first patent to be granted in the name of an AI. Ryan wrote to me about this latest development and said they had a ruling in their favor two days later in Australia as well. "The consequences are that this is the first jurisdiction to allow an AI-generated invention made without a traditional human inventor to be protected. This means that the system will appropriately encourage people to make and use AI to generate socially valuable innovation, rather than restricting the use of AI in R&D. It also means that, at present, an AI-generated invention can only be protected in certain jurisdictions and not others, which presents a real challenge for global commerce and innovations not restricted to national borders."

For more discussion about this, go to episodes 50 and 51.

In today's news ripped from the headlines of AI, we've got research with amazing and disturbing implications. We're going to spend a while on this. In their arxiv.org paper, 22 researchers from 15 institutions in 4 countries showed how machine learning trained on chest X-Rays, mammograms, limb X-Rays, and cervical spine X-rays could predict the race – white, black, or Asian – from those images with usually much better than 90% accuracy. This is not something a human radiologist can do.

They then added high-pass filters to the images and trained the models again, and the models maintained performance well beyond the point that the "degraded images contained no recognizable structures; to the human co-authors and radiologists it was not even clear that the image was an x-ray at all." If you look at the paper those images look like plain gray rectangles; and the AI was still able to figure out the race. I'm going to quote them some more because it's not necessary to add interpretation or drama to what they said; they spell it out in the paper. "These findings suggest that not only is racial identity trivially learned by AI models, but that it appears likely that it will be remarkably difficult to debias these systems. We could only reduce the ability of the models to detect race with extreme degradation of the image quality, to the level where we would expect task performance to also be severely impaired and often well beyond that point that the images are undiagnosable for a human radiologist. Overall, we were unable to isolate image features that are responsible for the recognition of racial identity in medical images, either by spatial location, in the frequency domain, or caused by common anatomic and phenotype confounders associated with racial identity." So one thing they did was get the models to isolate the areas of the images where the racial information was strongest and block out those areas; didn't help.

So why is this important? Well, medical institutions want to be able to process patient data without racial identity so as to avoid bias. This paper shows how hard that is. "One commonly proposed method to mitigate bias is through the selective removal of features that encode protected attributes such as racial identity, while retaining as much information useful for the clinical task as possible, in effect making the machine learning models 'colorblind'. While this approach has already been criticized as being ineffective in some circumstances, our work further suggests that such an approach may not succeed in medical imaging simply for the fact that racial identity information appears to be incredibly difficult to isolate. We strongly recommend that all developers, regulators, and users who are involved with medical image analysis consider the use of deep learning models with extreme caution. In the setting of X-ray and CT imaging data, patient racial identity is readily learnable from the image data alone, generalizes to new settings, and may provide a direct mechanism to perpetuate or even worsen the racial disparities that exist in current medical practice."

Wow. What can I add to that? It just goes to show how many surprises AI can spring on us, and we're still in the very early stages with it.

Next week, I'll be talking with Amit Gupta, who is the founder of Sudowrite, an AI-powered creative writing app that I got to try, and which produced some riveting and astounding results. 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