AI and You

Transcript

Guest: Kordel France, part 2

Episode 73

First Aired: Monday, November 8, 2021

Hello, and welcome to episode 73! Today we're going to conclude the interview with Kordel France, founder and CEO of Seekar Technologies, and AI engineer. Seekar is currently building the first clinical AI tool used to advise neuropsychologists in diagnosing mental disorders. Last week we talked about an amazing range of topics from autonomous tractors to autonomous weapons. This week we'll be talking about explainability, artificial general intelligence, and the fragility of image recognition AI, among other things. Let's get on with it.

What's in the pipeline next? You've covered a lot of territory here.

So something we're particularly interested in is really trying to push AI more towards the artificial general intelligence realm, which is basically trying to get out of these really narrow scopes of application and broaden the horizons a little bit to what AI can do. So one thing I mean by that is, we're working on an effort right now where we have a robotic-- So we're using our machine vision platform, that, and we're also using an audio processing artificial intelligence module. And these two, you can imagine a robot that can basically see and hear. And the robot gets damaged, and its audio processing module gets damaged, so it can no longer hear, but it can actually use the software that that processing module was built upon and recruit that artificial intelligence module to be used for vision, so your vision sharpens. So why is this important? If a human gets hearing damage, but that part of their brain is not damaged, let's just say for some reason, there's a significant event, there's a traumatic event that ruins your hearing. And when you can no longer hear, that part of your brain will actually start to kind of shut off because it's not inputting any sound. But your visual cortex will actually start recruiting resources over and start using that part of your brain that was used for audio processing to compensate in vision. So your vision in some regard becomes a little bit sharper. And there's different ways in which this happens but we're trying to simulate the same thing in robotics. So a robot becomes damaged. Can it still recruit any resources from what was damaged to help sharpen its other senses, essentially? And this is obviously a big advantage to robotics, but it's also a big help in medicine as well, because if we can try to get to a point where we can simulate parts of damage to the human brain, then we're able to provide a playground for physicians to go into and run simulations to say, "Okay, I can better understand how this condition manifests because I've just run a simulation through Seekar's software." We're a long ways away from the latter portion because the human brain is so complex and to accurately emulated it is going to take quite a bit of time. But that's the goal in which we're trying to develop this tool for.

I've not met many people that are covering so much different territory. How big is your company?

So right now, we're only at about 11 people. We've come a long ways, and we move very fast. We're actively recruiting to hiring more people because we're getting ready to raise a Series A. But yeah, we moved very fast with the pandemic. It was very good to us, and this medical technology got a lot of attention from physicians.

So you mentioned artificial general intelligence, that is a loaded term, a high bar to jump, and, obviously, a lot of money chasing a goal that some people say is either impossible or hundreds of years away. Talk to me about where that bar is for you and why you use that term.

The bar for me is still… I don't know if it's hundreds of years away, but it's definitely at least towards 100 years away, in my opinion. If not 100, at least towards the upper end, or the north end of a century. And the reason is because I think our definition of AI starts to shift with time. AI back in the '90s was basically clustering and regression, very sophisticated mathematical regression. And now you say that that's artificial intelligence and you might get laughed out a little bit. So our threshold of AI seems to kind of move a little bit, but we're expanding the capabilities of it. But we're getting very, very good at building these very narrow applications of AI but coupling them all together is very hard. So when you have a computer vision platform that can identify objects, and you have an audio processing platform or artificial intelligence model that can identify sounds, and you couple them together, it still isn't even close to the point where it can say, "I am. I know who I am. I know where I am within society. I am conscious." So being able to replicate a small minute portion of consciousness, I still think it's probably at least 50 years away, if not even more.

Well, let's parse some distinctions here. Consciousness doesn't necessarily have to be present for artificial general intelligence if you're talking about the capability to solve general problems.

That's true.

Although I think that you could argue that it might turn out to be necessary because we still don't know how to do it. Then you've been talking about some advanced capabilities with AI that move in the direction of AGI, and I think that AGI may be something that we creep up on. One of the things that we're discovering about artificial intelligence right now is that even the narrow ones are able to do things that we would previously have thought you'd have to have AGI to do, like some of the things GPT-3 is doing. So I suspect it's going to be this rolling, creeping thing where one day we've kind of sort of passed the Turing test and the next day, yeah, it's better and it's fooling more people. And not that that's even the best test now for that kind of thing, but that it's going to come upon us incrementally. So you're working on something that you labeled as heading in that direction; you're maybe not going to wait 100 years for that. How do you see it advancing incrementally before then?

So you bring up the Turing test, which is a good point to kind of help augment my point. And I think the direction in which we can head is really trying to satisfy the Turing test for different senses. So I can have a conversation with a smart home assistant, and vocally that, I mean, in some regard, it passes the Turing test, right? You can't have a deep conversation with your smart home assistant, but you can ask it pretty decent questions in which it will return pretty

sophisticated, pretty decent answers as far as Google's search queries go. But being able to take that to a level of machine intelligence or machine vision, right, so we have what's considered briefly that that satisfies some aspect of the Turing test and audio processing. Machine vision seems to be satisfied in some respects with the Turing test in self-driving cars because cameras are getting pretty capable of identifying everything around them. But they still need to take that leap to be able to identify many more things above that. An autonomous car can't navigate through a house, it can't navigate up the stairs, right? Because one, it's mechanically not capable. So part of it's a mechanical problem as well. And being able to solve the Turing test in these very narrow scopes, I think will collectively allow us to bridge an overall larger Turing test that we can satisfy by enveloping the narrow technologies together. But yeah, if we can satisfy the Turing test in very narrow scopes, I think that will collectively help pose AGI a lot faster.

Right. If you're having a conversation with your thermostat, it's a lot easier for you to imagine that it's artificial general intelligence when the conversation is restricted to how warm it is in the living room.

Exactly.

And you talk about autonomous vehicles being able to recognize things around them, but that they can't go upstairs. What do you think about Tesla's announcement of their robot that they are developing which would be able to go upstairs and is powered by the same hardware as the Tesla car?

I've thought a lot about this because I watched their presentation when it first came out. And their approach to computer vision that they're doing with their cars, they completely removed radar. I mean, they've got a massive amount of training data that they can train this robot on, so it's not like they're training it from scratch. They're going to be able to take a lot of the data that they've learned from their cars and navigation to some respect and be able to really just put that into a humanoid robot. Being able to release it by the end of 2022 I think is what they promised, I think that's a pretty lofty goal. I don't think that's going to happen. And primarily I think it's a mechanical problem. I mean, it's very hard to get a robot just to balance and walk and to be able to do-- I mean, I don't know if you've seen the Boston Dynamics videos, but those things are incredible and that's a pretty difficult task. I mean, it's doing some things that I can't do. And now Tesla is saying it's going to be able to do all of that and going to be able to do tasks around the house and have a casual conversation with you or a specific conversation with you. I definitely think it's possible, but I don't think it's going to be to the point where it will pass the Turing test. I think it'll basically be a smart home assistant with a very sophisticated mechanical body. And I don't think it's going to be next year, but they've definitely got the infrastructure as far as artificial intelligence training goes to build it.

I would agree. They haven't delivered autonomous driving yet, and a humanoid body navigating in arbitrary environments is I don't know how many orders of magnitude harder a task. It may be somewhat safer in that it's not weighing two tons and moving at 70 miles an hour. So it might be a less regulated environment, one that's harder to make fatal mistakes in but nevertheless-- Well anyway, we can fill in those blanks. So you've got some thinking going on

here obviously about higher-order things, artificial general intelligence, and so forth. So tell us your philosophy on artificial intelligence as threat versus utopian boon. Where do you think we're going ultimately with this?

I am a part of the camp where I think artificial intelligence could be an existential threat. However, I don't think we're going to see it anytime soon. And I think the threats more pertinent to our generation now are not in the form of Terminator or iRobot, I think they're more of the form of all of the… just to keep using Tesla again, all the Tesla vehicles are hacked, and they're all instructed to drive off the east coast, right? That's a pretty big existential threat, being able to have all these robot cars act in, you know. And that's an AI existential threat in a less humanoid form, but still one that is pretty powerful. Or you have an artificial intelligence-based algorithmic trading fund that manages a very large portfolio, and it liquidates everyone's accounts and goes rogue, and partitions the money in ways in which it wasn't intended. Those uses of AI I think are more pertinent and things that we can try to reconcile now because those are incremental steps to being able to actually build a more existential humanoid threat such as Terminator and iRobot, etc., things that we see in the media. But if we can try to protect against malicious AI in that regard now, and to the development of our technologies now, I think that that will-- I guess what I'm saying is we need to start having those conversations now because, by the time it is a threat, it's going to be too late. We need to start being able to have a safety harness in place for some of these things that are being developed in the future. So I think we need to start thinking about what that looks like, and how we can start making sure ethical AI is being developed, how we can make sure that bias is rid, and we see this a lot in facial recognition, how we can make sure that bias is removed in those regards, because if we can imagine that on a much higher magnitude, bias in a Terminator robot is going to have a much larger effect in hundreds of years than it's going to have in a camera system now, and it already has a bad effect. So we need to start looking at that, and mitigating a lot of bias in my opinion.

Yes. One can imagine all kinds of other threats, like what if the AI in the hunting scope gets a bit flipped and now it decides that the image of another hunter is a valid target.

Exactly yeah, no, that's true.

Let's pull it back to today. You've got this perspective on AI in many applications so let's use that, let's tap that for a moment. Where do you think that AI with this huge range of applicability, its flexibility, where is it not being exploited yet, not being used, that you see openings for, or where it could or should be used more than has happened yet?

I see a lot of over-promise and under-delivering with AI. So a lot of 200x improvements, 300x improvements, that type of thing are promises. I think we need to start trying to hit the five to 10 to 15% incremental improvements and try to augment people's capabilities. So what I always say in regards to medicine is we're trying to help physicians do their job 15% better with artificial intelligence, 10% better with artificial intelligence, and then we'll worry about making them twice as efficient, three times as efficient, etc., and really trying to augment their capabilities of what they can do, augment their current skill set. Because at least from the people we've talked to, a lot of the apprehension from being used with some sensitive applications, such as defense

and medicine, is that a lot of the narrative is that we're trying to replace people, or we're trying to replace jobs. I don't necessarily think we should take that angle. We should try to help people out and make us hyper-productive, incredibly efficient at our jobs so that we can just make technological leaps with the help of AI, not in lieu of it or in spite of it. But that's one thing is kind of maybe trying to take a more augmenting approach versus replacing approach. The second thing is explainability. A lot of these AI systems are incredibly intelligent, and they do deliver robust results until they don't. And when they don't, some of them can't be explained, because a lot of them have this black-box approach. So inputs go in, and it outputs a result and outputs a classification, but you can't really explain how all that happens on the way in. If we can build algorithms that explain themselves more or less in a fashion that can be derived in the sense that a human can explain themselves, then I think that that's very key to moving forward. I mean, if you can imagine if GPT-3 could explain how it devolved to a certain solution, that's quite significant, right? And that really makes an order of magnitude leap in artificial intelligence progress as far as passing Turing tests in some regard, and really just a big leap forward towards AGI. But making simple neural networks such as object recognition systems explainable, you know, why did I misclassify this? Why did this facial recognition system misclassify this group of people? Well, let's go back. Was it the training data? Was it the algorithm? What was it being able to kind of hone in on why it misbehaves is pretty key to making sure that we make sure AI is moved forward in the right direction?

Right. And those explanations though, usually come from the people. Like an AI often misclassifies as a picture of an empty landscape as sheep because every picture of a sheep it's seen they're in a field. So its idea of sheep also incorporates the idea of field and the white fluffy things are kind of optional in that. But it took humans to figure that out. Now, I can see how people would want an explanation of how your app decided that someone had a certain diagnosis, medical condition from the way they sounded, the way they looked. But telling them, "Well, we've got these 10,000 weights in the neural network that produced that answer" doesn't do any good. So is there an explanation available?

So that's one thing that we try to differentiate ourselves as a company with is being able to provide explainability with our solutions. So we actually have two different types of explainability reports. One is for the engineers that actually get a neural pathway, if you will, of all the neurons that were fired into the neural network so that they can go back and see if there's 100 paths that are being consistently taken and there's 1000 others that are being avoided, why is that? Obviously, that's a representation of an unbiased data set or something wrong with the algorithm. So that's a very technical explainability report. But we also in that regard, built another report for physicians, or our users that allows them to say, "Okay, the algorithm thinks that this is COVID-19 because it identified this, this and this, and these different biomarkers." And we do this in a couple different ways. But being able to say, "I thought that this pattern by the sternum of the left ventricle of the lung in the extra image was indicative of COVID and so that's why I classified this image as COVID-19," that's pretty helpful. And the doctor can go back and say, "No, why are you looking there? There was never any patterns of COVID there." But these types of conversations should be facilitated by our artificial intelligence systems, and that's something we're trying to do with Seekar. And a lot of the apprehension that we've seen

with certain industries being able to use AI is because it's not explainable, and so we're trying to help kind of cater to that camp.

Now, I recall seeing images that degraded by a certain specific kind of noise, which then wrecks the machine vision algorithm - to give an example, a picture of a giraffe - and the researchers add what looks like invisible noise to it to a person and now the algorithm is more sure that it's a gibbon than it was that it was a giraffe before that. And that clearly violates the premise that the AI is recognizing image the way that people do, it's doing something very different. I assume you're familiar with what I'm talking about?

Absolutely.

Now, that would then suggest that the reason reasons that identified it as a giraffe are not related to the way that we would identify it as a giraffe, they're related to some kind of, to me, incomprehensible patterns of pixels that were picked up in the training data somehow and renders it inherently fragile. Have we advanced beyond that? Are we more robust now in our image recognition?

So I've thought a lot about this in regards to a lot of the autonomous vehicle companies because they've got to accommodate quite a large environment, and being able to identify a car in a snowy environment versus the rain, versus the fog, versus a freeway, versus a dirt road, I mean, these are very drastic scenarios. And I would hope that adding a couple pixels of noise to one of their camera feeds is not going to drastically increase or actually increase at all their noise.

But let me be clear about this. The noise that was added was not random noise, it was noise picked by the researchers pulling it out of the neural network so that they knew which pixels in the image were key to the algorithm so they could add strategic - what looked to us like noise but strategic - pixel changes that were invisible to us but still radically changed the machine interpretation.

I see, yes. And I see that a lot of that work, at least from what I read was attributed to an adversarial attacking and how do you fool an image recognition system kind of a thing. But yeah, that's a great point of things that we need to reconcile. So have we even solved computer vision yet? If the model is not learning what we think it's learning, then we need to go back to the drawing board and really re-address the algorithms that we're using, to begin with. And you imagine the effectiveness of that on something as catastrophic as a facial recognition system. Imagine if someone were to exploit that system in the ways that you just alluded to, that's terrible. And that's something that we need to be addressing,

Well, it's already been done. I think you can buy t-shirts that foil facial recognition algorithms. There are examples of signs where it says stop, and they've added pieces of tape to it that don't even cross the letters that change the interpretation to a 45 mile an hour speed limit sign. And, of course, they knew exactly what they were doing and how to do that. But it did suggest to me that if you couldn't ask that network, "Why did you interpret this as either a stop sign or a 45

mile an hour speed limit sign?" and expect to get back any kind of answer that related to the way we interpret those images.

Yeah, no. And that's where that explainability module comes in because we need to be able to know what patterns are you recognizing speaking to a neural network? What patterns are being recognized? What do you think are indicative of a stop sign? Show me. And maybe you have a visual representation of being able to highlight what they are. And there are some efforts that are undergoing this like, I think with faster RCN object detection, but that's particularly where this explainability feature is needed. I agree.

Yeah, fascinating. Well, it's been an amazing conversation. What are your long-term views? Let's say 10 years from now, what would you like to be working on in AI? And where do you think we'll have advanced to?

So my career goal is I want to be some contributor to being able to advance AI to the point where I could look back on my life and say I was a key contributor to AGI, or simulation of some aspect of consciousness. And I think we're on that realm where we are now with some of the projects we're working on that I just spoke about. And if I can look back at my life and say that I've done that and I've contributed heavily in those regards, then I'll be happy. But I do think that AI as a field in general, we should focus on really making sure that the narrow approaches we do have are extremely robust against a lot of adversarial attacks that we just spoke about before we move on to even more sophisticated AGI techniques and more sophisticated AGI products. So I hope that we can kind of remove the sales hype and really ground ourselves to make sure that we understand what we're building and even though we can build it, should we build it kind of a thing.

What you're talking about there, the whole ethical view there, is in the language of defense, a target-rich environment as far as AI goes right now. How would you like people to follow you and your thoughts and what you're up to?

Yeah. So on our website, seekartech.com, we offer a free consultation just to discuss ideas or approaches. We're trying to democratize AI and address new aspects, new realms of application for artificial intelligence. So personally, I always hop on those calls just to verify that we can or don't have the capability. We've got a newsletter you can subscribe to through our website as well. And then you can find me on LinkedIn or Twitter, which I post all things AI under kordelkfrance on both platforms.

Terrific. Thank you, Kordel France, we've learned a lot here. I'm going to pay a lot more attention next time I drive by a farm and also, maybe next time I'm anywhere near a duck hunt. But thank you very much for coming on the show.

Peter, thank you so much. I'm a big fan of the show. Thank you for your listeners too, as well. I really appreciate your time today.

That's the end of the interview. I hope you were as riveted by it as I was. We've got enough listeners now that I know it's not just me that wants to know this stuff. I've not talked about myself a whole lot

on this show, because - it's not about me. I remember when I first engaged a publicist after writing my first book and they were setting me up for interviews and the chief guy – wonderful man called Mike Onorato – said, "They interview the author, not the book." In other words, interviews would ask about me, my story, my history, my life, before they talked about my book. My first thought was, "Oh no, I've made a huge mistake. I'm not interesting." My second thought was, "Why would people want to waste their time hearing about me when I want to tell them about how the world might end?" But that was the reality, and I learned to deal with it, and if you've listened to me talk enough you've heard about me coming to terms with Asperger's Syndrome enough to be able to do TEDx talks and speak to a British parliamentary committee, which I would never have dreamed of before this, but because I have kids and I want to do the best I can for their future, I was highly motivated – actually more than that, I was called – to go beyond my previous self. And I don't tell people about that as a way of tooting my own horn but to show that if you believe in something enough to want to make a difference, you too can go beyond what you thought you were capable of. And we need that. It turned out that I didn't have to wait to be ready, whatever that meant, but to me it was some vague notion of being completely comfortable on a stage and not an introvert any longer and being gregarious and – well, a whole lot of things that it turned out I didn't need to become. In fact I was already ready and I didn't know it. What have you been waiting to be ready for?

Some people I know think this is what I should talk about all the time, the personal story telling. That's not going to happen. I'm not saying that personal storytelling isn't heartwarming and inspirational and all that, but it's not in short supply. If that's what you want, turn on the Oprah Winfrey Network – assuming that still exists, for all I know it doesn't, but in that case turn on the Lifetime Network or the Hallmark Channel and you can see that kind of thing 24 hours a day without even having to change the station. But media that helps you get a practical, useful understanding of the most revolutionary invention since electricity, that makes it as accessible as possible to the greatest number of people, that teases out the nuances of what could be the human race's ultimate destiny or final exit, well, there's not enough of that. So there's one of our periodic gut checks about why I'm doing this show.

In today's headlines ripped from the news about AI, there is a new robot in space, and this one is talking with astronauts on the International Space Station. Cimon is like a white basketball that floats around the ISS and can answer questions and access experiment and mission data. Actually Cimon has visited the ISS for the first time a few years ago, but now its successor, CIMON-2, is going to the ISS this fall with astronaut Matthias Maurer to be part of a study about AI astronaut assistants. Cimon is called a "floating brain" and is about as good at natural language processing as Siri, but this was developed by the German Space Agency in collaboration with Airbus and IBM as an experimental assistant for astronauts. In a video from ESA you can see German astronaut Alexander Gerst talking with it and telling it to move around. The response time is instant despite the fact that the processing is done in the IBM Cloud in Frankfurt and the signals have to travel down to Germany and then to the Lucerne University in Switzerland on their way there. It has a screen which displays a simple face most of the time. On the video you can see that when it doesn't understand something it admonishes Gerst "don't be so mean" and behind him you can see American astronaut Serena Auñón-Chancellor giggling.

Next week I'll be talking with Michael Hind, who is a Distinguished Research Staff Member in the IBM Research AI department in New York and we are going to talk about explainability. I wanted to find a top expert on that subject, Google said it was Michael Hind, and he agreed to come on the show. You've got to hear what he said. 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