EPISODE 103

[INTRODUCTION]

[0:00:10.8] SC: Hello and welcome to another episode of TWiML Talk, the podcast where I interview interesting people doing interesting things in machine learning and artificial intelligence. I'm your host, Sam Charrington.

Contest alert. This week we have a jam-packed intro, including a new contest we're launching, so please bear with me. You don't want to miss this one. First, a bit about this week's shows; as you may know, I spent a few days at CES earlier this month. While there, I spoke with a bunch of folks applying AI in the consumer electronics industry and I'm including you in those conversations via these series of shows.

Stay tuned as we explore some of the very cool ways that machine learning and AI are being used to enhance our everyday lives. This includes work being done at Anki, who built Cozmo, the cutest little computer vision powered robot. Lighthouse, who's smart home security camera combines 3-D sensing with deep learning and NLP. Intel, who is using the single-shot multi-box image detection algorithm to personalize video feeds for the Ferrari Challenge North America.

Firstbeat, a company whose machine learning algorithms analyze your heartbeat data to provide personalized insights into stress, exercise and sleep patterns. Reality AI and Koito, who have partnered to bring machine learning based adaptive driving beams, or automatically adjusting high beams to the US. Last but not least, Aerial.ai who applies sophisticated analytics to Wi-Fi signals to enable some really interesting home automation and healthcare applications.

Now as if six amazing interviews wasn't enough, a few of these companies have been so kind as to provide us with products for you, the TWiML community. In keeping with the theme of the series, our contest will be a little different this time. To enter, we want to hear from you about the role AI is playing in your home and personal life and where you see it going.

Just head on over to twimlai.com/myaicontest. Fire up your webcam or smartphone camera and tell us your story in two minutes or less. We'll post the videos to YouTube and the video with the most likes wins their choice of great prizes, including an Anki Cozmo, a Lighthouse smart home

camera and more. Submissions will be taken until February 11th and voting will remain open until February 18th. Good luck.

Before we dive into today's show, I like to thank our friends at Intel AI for their continued support of this podcast. Intel was extremely active at this year's CES, with a bunch of AI, autonomous driving and VR-related announcements.

One of the more interesting partnerships they announced was a collaboration with the Ferrari Challenge North America race series. Along with the folks at Ferrari Challenge, Intel AI aspires to make the race viewing experience more personalized, by using deep computer vision to detect and monitor individual race cars via camera feeds, and allow viewers to choose the specific car's feeds that they'd like to watch.

Look for my conversation with Intel's Andy Keller and Emile Chin-Dickey later in the series for an in-depth discussion about this project. Be sure to visit ai.intel.com, where you'll find Andy's technical blog post on the topic.

Now a bit about today's show; in this episode, I sit down with Alex Teichman, CEO and Co-founder of Lighthouse; a company taking a new approach to the end-home smart security camera. Alex and I dig into what exactly the Lighthouse product is and all of the interesting stuff inside, including its combination of 3-D sensing, computer vision and natural language processing.

We also talk about Alex's process for building the Lighthouse network architecture, the tech stack their product is based on and some things that surprised him in their efforts to get AI into a consumer product.

Now, on to the show.

[INTERVIEW]

[0:04:27.0] SC: All right, everyone. I am here at CES and I've got the pleasure of being seated with Alex Teichman. Alex is the CEO and Co-founder of Lighthouse. Alex, Welcome to This Week in Machine Learning and AI.

[0:04:37.1] AT: Excellent. Thank you.

[0:04:38.6] SC: Absolutely. Great to have you on the show. Why don't we get started by having you tell us a little bit about your background? You've done some interesting things in the AI sphere.

[0:04:50.1] AT: Thanks. Yeah. My background is in perception systems for self-driving cars. This is all about getting them to understand what they see in the world, what is a car and what is a bicycle, what is a pedestrian and that sort of thing. I joined Sebastian Thrun's lab back in 2007 right when the DARPA challenges were wrapping up.

[0:05:14.9] SC: What were some of the specific things you were working on there?

[0:05:17.8] AT: My focus was on how you use 3-D sensing, Lidar in particular in that case, to do a better job of understanding what you're seeing in the world, you being a self-driving car or a computer more generally.

This is very different from using a regular colored camera to understand what you see in the world. When you have a 3-D sensor, you've got the full structure to work with in real-time. That opens up a variety of different computer vision techniques and it makes many of the very difficult computer vision sub-problems quite easy. Not all of them, but it makes many of them easy.

[0:05:54.5] SC: Can you give us an example of that?

[0:05:55.9] AT: Yeah. For example, segmentation and tracking of objects for which you have no computer vision model is extraordinarily difficult with regular video. When you have a 3-D sensor and when certain assumptions are met, then you could do a very good job segmenting and tracking objects, even if you have no idea what they are, if you want the semantic information and whatsoever.

This is something that's made use of very heavily in the self-driving car world, where you can see that there is a physical thing in the structure in the environment and it's moving around. You don't have to know what it is to drive safely around it.

[0:06:30.7] SC: Yeah, there seems to be a – there's a school of thought in and around the self-driving cars that taking advantage of what you're describing using Lidar and things like that, but then there's another school of thought where folks are saying Lidar is too expensive to be on every production vehicle. We're going to try and do things with just cameras. Any thoughts on that?

[0:06:55.7] AT: It's hard.

[0:06:57.3] SC: Which one? All of it is hard, or –

[0:06:58.8] AT: No, no, no. Well, getting all the stuff to work with just regular colored cameras, it will eventually happen. The information is all there and humans do it with what essentially amounts to just a camera. Stereo is not effectively – well, is not very effective at that range. We are just machines in some sense, very complex, very sophisticated machines, but we are able to do it.

The information is there and eventually we will get computers to be able to do that thing. We seem to be a long way off from that. It is quite hard. This is why virtually every self-driving car project is using Lidar, because it makes many of those hard problems a lot easier.

[0:07:39.7] SC: Fast forward to Lighthouse, what's Lighthouse up to?

[0:07:43.8] AT: Yeah. The story of Lighthouse. We're talking a lot about self-driving cars here. What we're doing is basically we're taking that set of computer vision and machine perception techniques and we're translating that from the self-driving car world into the home. That's the technology perspective on what Lighthouse is. That's the machine learning perspective.

From the customer perspective, Lighthouse is – imagine you had a traditional home camera, but it had the intelligence of something like Alexa or Google Home. It's a new interactive assistant that's based on this 3-D sensing and computer vision and cameras that lets you tell it what you care about, and then it tells you when it sees those things happen.

[0:08:30.2] SC: Then is the application – is it security, or personal virtual assistant, or something beyond?

[0:08:39.0] AT: It's both.

[0:08:40.2] SC: Okay. Give me an example of how I might use it.

[0:08:45.5] AT: Yeah. One thing you could do with Lighthouse is you can say, tell me if you don't see the kids by 4 PM on weekdays. You literally just say those words, that's it. It understands what you're asking for. It has a very good computer vision model for what children and are and knows what they look like, and it knows that you're asking for by 4 PM on weekdays, Monday through Friday. If doesn't see children by that time Monday to Friday, it will send you a notification. If it does, then it won't bother you.

[0:09:14.3] SC: interesting. Few years ago, I had some crazy project that I was going to do around the house. One of the first things I started trying to figure out was presence. This was pre deep learning, CNNs, all that stuff. I started looking at NFC and all these other things, and it's just so obvious now that vision is – cameras and vision is the way to do this. What are some of the challenges associated with deploying a vision appliance, I guess and the home environment?

[0:09:51.1] AT: Well, everything in computer vision is hard to some degree, because that's new – we're just at the dawn of artificial intelligence here and all of these different techniques are very cutting-edge.

Really pushing the boundaries in what's possible with deep learning and combining that intelligently with the sorts of techniques you could use with 3-D sensing, and particular around segmentation and tracking.

There is a lot of complexity and a lot of difficulty around building the hardware to do this too. Because this is the first 3-D sensor that has 95 degree diagonal field of view, that can see how to – depends on the details, but 7 to 10 meters is typical. It's quite challenging to put all that stuff together. Hardware is hard, is the phrase for a reason.

[0:10:42.1] SC: Is the device itself, is it a connect plus the camera?

[0:10:47.0] AT: You can actually think of it that way. It uses a different underlying depth-sensing technique. Actually it depends on which connects you're referring to. The original connect was

structured light and it's like stereo. There is a texture pattern going out. A projector and you know where that projector is and there's an infrared camera and you can try and deal it from that. That was the original connect. It was actually what we prototyped Lighthouse on in the very beginning.

What we're using now is a ton of flight camera, where that sends out modulated light and then you look at the phase shift between that modulated light as it returns on a reference signal and that phase shift tells you how far away things are essentially. Every pixel in the image, not only do you see, "Oh, it's like this shade of brown." You also see it's 3.72 meters away. You can get that for the whole scene.

[0:11:39.9] SC: Interesting. I can't remember the name of this thing. That was a Kickstarter that I backed. I haven't done anything with this thing yet, but it was like a mini-Lidar scans, I think was the name of it. Have you ever come across that?

[0:11:50.7] AT: I haven't come across that one, but was it for scanning your face, or yourself?

[0:11:54.8] SC: No. It was like, it was for hobbyists. You could put it on a mobile robot and just experiment with it, that kind of thing. I don't think there is any specific end-user or use case associated with it. It was focused on – I think it was just an example of how to scale Lidar down to something if it's in the palm of your hand and is relatively cheap.

[0:12:19.4] AT: I see. I'm not familiar with that one, but both generations of the connect are a good example, the iPhone X with the 3-D sensor built into it and that's also a good example. They use that to make face ID actually reliable. Then self-driving cars obviously with all the different varieties of Lidar that's out there.

[0:12:38.9] SC: Maybe a little bit more detail around the some more examples of used cases for the device itself might be helpful.

[0:12:48.0] AT: Yeah. I mentioned the one about, like if the children don't come home by a particular time.

[0:12:52.3] SC: Is it just children, or is it like if Bobby doesn't come home, or Susie doesn't come home. Are you able to identify specific faces and associate them with kids, or is it just children?

[0:13:04.8] AT: Yeah. You can do either in fact with Lighthouse. Lighthouse has the ability to understand, "Oh, that is a child generally." It also has the ability to understand faces of specific people, and in particular what that is most useful for is – you can do something like say to Lighthouse, "Hey, tell me if you see someone you don't recognize while Cindy and I are away," for example.

That lets you get at – I don't know if your children bring home a new friend while you and your wife are out at work or something, and you might just want to know like who is this new person and it will tell you about it. It will send you a push notification when it sees that.

Or if you have a dog walker, or a babysitter and one day it's somebody different, or somebody new is there. It will proactively notify you about this. You don't have to go back and check every day, because you have setup this alert with natural language.

[0:13:57.0] SC: Yeah, it was just occurring to me that as you were describing these use cases that as complex as the computer vision and the 3-D sensing is there's also an NLP challenge, like how do you capture the full breadth of what someone is going to want to ask this thing. We've talked a bit on the podcast about some of the underlying NLP technologies and spoke with someone on the Alexa team. Are there unique challenges associated with the way you're using NLP in the context of this device?

[0:14:34.5] AT: I wouldn't say there is necessarily, I don't know, unique research challenges on the natural language processing side. There are difficult and important engineering challenges that we need to know on that side of things. It's the computer vision where the really heavy-duty research grade techniques are being deployed, at least for the current generation of Lighthouse. I mean, you can imagine Google Assistant needs to answer virtually any question you could throw it.

With Lighthouse, there is actually a more restricted set of things. If you ask Lighthouse for directions from here to wherever, like we don't do that. That's not what we do. For the set of

things that we understand on the perception side, like we're – actually, we're very good at being able to answer those questions. It's a more constrained space, that makes the problem easier, there's more structure in it.

[0:15:27.5] SC: Okay. When you're providing the user manual for this thing, are you telling someone these are the 10 things you could ask it? Or are you setting the expectation that they should just be able to ask it things related to the kinds of stuff that it can do?

[0:15:43.2] AT: Yeah. We float rotating suggestions in front of people in the app. When you're in the natural language interface screen, you'll see here is the set of things, you might consider asking some examples of these things in this category and guide you through what categories of things we understand.

That includes, for object recognition it's people and children and pets and that kind of thing. For action recognition, we recognize waving at the device. You can say something like, "Hey, tell me if you see someone waving, hello while I'm out," that kind of thing. We understand time ranges. We allow you to setup alerts for things that happen in the future and so on. We'll guide you through those different categories of what we do.

[0:16:31.8] SC: On the computer vision side, what are the key research level challenges that you're tackling?

[0:16:40.5] AT: What it is really coming down to is applying deep learning at a large scale with 3-D sensors, combined with colored cameras. There are particular things that this setup let you do, that you just can't do in any other domain. For example, the 3-D sensor lets you segment and track objects through the space, without you having to have any semantic understanding. You don't have to know what that thing is. You just know it's a thing and it's moving through the space.

Now, your unit of classification from a deep learning point of view is that segmented object track through space and time. This enables several things. One, it's just more accurate, because you have more views of an object as it's moving about and you can integrate all of that information. Two, to very, very natural setup for doing action recognition, because you've got this thing moving through space and time and you can ask questions like, "Is this a dog? But also, is this a

dog jumping up on my couch? Or is this a person waving hello?" And so on. It's a great setup for working on these kinds of very challenging computer vision problems.

[0:17:47.5] SC: Okay. You talked about segmenting these objects and I'm thinking about this primarily being driven by the 3-D sensor. In what ways does having the camera augment what you're able to do beyond just the 3-D point law?

[0:18:03.0] AT: Well, so at deep learning time, it's a specialized architecture that's using both of those channels. Almost the attentional mechanism if you want to call it that, that's primarily driven by the 3-D sensor. But then once you're analyzing what is this thing, now use everything we have and that is including the 3-D sensor data, the point cloud of the objects, as well as the colored camera data. We combine these things, deep learning architecture that uses both of those and then merges them, and then goes into an LSTM for doing understanding of what is happening over time.

[0:18:44.0] SC: Okay. What was the process for coming up with the network architecture for this thing? Have you thought of something off the shelf like inception, or name your network architecture, or did you build it up from the ground up?

[0:19:01.8] AT: Yeah. I mean, in this context it always makes sense to start from a baseline that's reasonably easy to just pull the thing out of the box into play and see what happens. We did that with Google Net a hundred years ago just to see what would occur. Yeah, it did something. It was good.

It was pretty clear that we need to customize this thing to get the level of accuracy that we really want. Then the process from there is, well are you familiar with the phrase graduate student descent?

[0:19:32.3] SC: Sure.

[0:19:33.0] AT: Okay. Yeah. It's intuition combined with significant perseverance combined with lots of compute.

[0:19:48.0] SC: I think the turn way of saying that post-NIPS 2017 is alchemy.

[0:19:52.4] AT: Yes. There is a lot of alchemy when it – I mean, it's sad actually in that a lot of my PhD work was during the age when proper machine learning techniques should be comebacks. Just like, "Yeah, it's a descent method. You're always going downhill and just roll to the bottom and you'll find the solution." It will be great.

Now it's just – it's not come back in just – maybe it's working and maybe it's not working. "Oh, I don't know. Try a different momentum term and maybe it will work this time." It has challenges and advantages too, like now things actually work. That's pretty cool.

[0:20:28.3] SC: For folks that are trying to productize around deep neural networks, what – in that way, I guess I struggle with the graduate student descent as the answer. I guess, probably we all do a little bit. Have you developed any intuition or rigor around, or methodology rather around the way you – the way you build out network architectures for this problem space? Or even maybe another question in this background is, was the network architecture like upfront work that you did and it's static, or is it – how rapidly does that evolve?

[0:21:09.2] AT: That is an ongoing effort in many different ways. In one way, we are collecting new annotated data all the time, both from our own early access testers who provide us access to their data for us to use for training purposes. Also, if there is a mistake in the field, you can annotate it as such, make use of it and improve the models. We have a stream of annotated data coming in, and so we're always taking the same network structure and taking that new training data and turning the crank and redeploying. That cycle – I mean, it depends on the details, but that's on the order of days.

The new architecture deployment cycle, that's more like weeks or months as we come up with some new idea of like, "Oh, maybe we can compress the network this way." Or maybe it would make a lot of sense to build out this piece of the network and then we'll go work very hard and validate that new network and find out, "Oh, indeed. This reduces compute time on our end and produces a better experience for the customer." Great. Let's go deploy this. It's all about large-scale quantitative testing.

[0:22:13.5] SC: You mentioned compressing the architecture. Are you deploying the network on the device, or are you doing inference in the cloud or something like that?

[0:22:22.1] AT: It's largely in the cloud at the moment. There is a variety of reasons that make sense. Although, I should mention it is not entirely in the cloud. It really is a distributed computer vision system to squeeze all the last bits of performance out of it that we can. You really do want it to not all run in one place. It makes sense to have some of it run on device, some run in the backend.

[0:22:42.7] SC: Talk a little bit about that in more detail. What is running on a device? How do you partition what's running on the device and what's running in the cloud?

[0:22:51.0] AT: Yeah. At a high-level, the device is doing the attentional mechanism. It's doing the segmentation and tracking of what is interesting and new. Then it is nuanced here, but at a high level it's doing that.

[0:23:04.4] SC: A simplistic perspective, you're not sending a bunch of frames up to the clod if there's nothing happening?

[0:23:12.3] AT: That's largely correct. Yeah. We do have to send some data once in a while; one frame every few seconds basically. This is actually so we can present to you a beautiful summary of the day. We call it a smart time lapse, or a daily recap, where you press one button and you get a 10-second or 1-minute summary of what happened during the day. It goes fast during the boarding parts and it goes slow when there's something of interest.

[0:23:39.4] SC: Interesting. Okay.

[0:23:41.2] AT: Yeah, generally we actually – we don't have to stream 30 frames per second, because it's actually not what customers really care about. Customers don't care about what were the RGB pixel values at like 3:47 AM yesterday. What they care about is, "Did my kids come home on time? What has the dog done since I left the house?" Because I just think it makes me feel warm and fuzzy inside. Was anybody new here and who was new here last week? Just show me these things." That' what they care about.

With traditional home cameras, we're a wash in data, but we don't have much useful information and that's what Lighthouse is all about is taking that enormous stream of data and compressing it down into just the bits that you actually care about.

[0:24:27.0] SC: Interesting. One thing that I'm curious about being here at CES and seeing tons of different consumer-oriented products that are trying to incorporate AI in one way or another, are there any things that you've learned that were surprising about pulling AI into consumer-oriented products?

[0:24:49.9] AT: Yes. Actually, when we started Lighthouse, myself and my co-founder, I thought it would be the AI problems that were the hardest across the board. They are hard for sure. There's no question of that. It turns out there is other hard problems that you have to solve along the way. For example, getting the UX right, like the UI and the interface and the user experience really right. That's actually quite difficult. It's something we spend a lot of time on, because what we – ultimately, the reason we exist is to deliver a delightful and useful experience to our customers.

We're able to do that with AI, but that's not the only thing. It can be quite hard to get those things right. Especially in breaking new ground in a new interactive assistant, how does one actually build the best interface to this thing? It takes a lot of work and iteration.

[0:25:46.2] SC: Do you have the Lighthouse laws of effective intelligent user experiences on? Are you boiled, you've learned down into key ideas that you tell a new team member?

[0:26:02.8] AT: I'm not sure we've refined it to that point, where I could concisely communicate something.

[0:26:08.5] SC: It's interesting. I've asked people this on and off for the last couple of years. I think that – it strikes me that we've developed a fair amount of methodology around traditional user experience via mobile, via the web and it strikes me that there is some set of rules that will evolve around designing intelligent systems. It's too broad, but presenting intelligent experiences to consumers. Haven't really found – no one said, "Yeah, I read this book about it." We're still too early for that.

[0:26:54.3] AT: Too early for those. There is one guiding principle actually that is worthy of mentioning here. It's always in the back of my mind with this kind of interface. The reason it exists is to make useful information accessible to you as quickly as possible. That's the reason natural language interfaces are good. Stepping outside of Lighthouse, looking at something like

Alexa or Google Home, one of the reasons they're so good is because you don't have to go find your phone or pull your phone out of your pocket and unlock it and go to this, go to the right app and then play your music and say, "No, play it on this interface." Then finally, it comes out. You just yell across the room, "Hey, play this thing." This works.

The reason that's amazing is because it saved you 10 seconds. It seems so trivial, but it's not. It really, really, really matters. When you look at this from the – I don't know if you want to call it the nerd point of view, so this is me, but it's all about reducing latency and increasing bandwidth in the human machine interface. That's the point of natural language is that you have a though in your mind. There's a thing you want to do. Right now, generally you have to translate that into, "Okay, I'm going to pull up in my phone. I'm going to tap on these buttons to get to the right app and I'm going to tap on some more buttons to do the thing I'm trying to do. I have to go to this menu and adjust the slider bar." It's terrible.

What you should do is there is that thought in your mind, just say the thought. That just happens. That is just so much better. That is an – I don't know. It's an order of magnitude improvement in latency in that interface between this intelligence in my head and this intelligence in my phone.

[0:28:38.5] SC: On the NLP side of things, did you start out with any of the popular cloud-based platforms for doing that stuff? I forget what it's called now. X.ai, you know what that is, but all the cloud vendors have their own. Or did you roll your own?

[0:28:56.9] AT: Right. They are useful prototyping platforms. There maybe even become applications where they get you all the way. That is not the case for Lighthouse. I mean, I can tell you that for sure, because I used one of them over a weekend to produce a little demo of like, "Hey, this is what I have in mind. I think this might be a way to really nail the user interface for this thing."

By the way, actually, I mean, when we started Lighthouse we knew the direction to go into to solve the perception problems, but we didn't know how to solve the UX problems. It was only along the way that we discovered that like, "Oh, my God." The natural language interfaces are the way to do this. It's not possible as far as we are aware to produce an interface with buttons and sliders and whatever else it might be to get you to be able to say, "Hey, tell me if you see

anyone new at the doorstep. I'll send in our way next week." How would you do that? You just can't. With natural language, it just works. It just works. It's amazing.

[0:29:55.6] SC: I've gotten that feedback quite a lot from folks that are trying to productize NLP like the platforms are interesting way to start, but you ran out of runway in terms of their flexibility and ability to get you all the way.

[0:30:11.8] AT: We build all our own. It's the only thing to do in this area.

[0:30:15.4] SC: Can you tell me a little bit about your tech stack generally?

[0:30:20.9] AT: Yeah, happy to. We use a lot of C++, because this is real-time.

[0:30:28.2] SC: On device and in cloud?

[0:30:29.1] AT: Both. It's real-time performance, memory-intensive computer vision running at scale. Well, either at scale on the backend, or on a limited compute device out on the frontend that is touching hardware. Both of these places C++ is the right thing to use at that level.

Now when we're prototyping a new architecture for our – the deep learning system, yeah it's totally reasonable to twiddle around in Python to have faster iterations on that, but ultimately when it's like building and deploying real systems, it ends up being C++.

[0:31:10.5] SC: Did you build out the NLP platform on C++ as well, like the whole system for intense and all that stuff?

[0:31:18.5] AT: I'm simplifying a bit, of course. The core computer vision system is in C++. There's a Java layer around that, because that's easier to interface with your phones for example. It turns out, that's also a good place to build your natural language processing. For whatever reason in academia and at least my circles of academia back in my Stanford days, natural language processing was generally done in Java. Computer vision on self-driving for example was all C++.

It probably is the case that on self-driving cars, C++ is a more natural fit because you have to interface with sensors and you have real-time requirements that's very heavy data. Whereas,

natural language processing is often less. In any case, that is a natural fit and our natural language system lives out there.

[0:32:04.1] SC: Is the Java ecosystem for the natural language stuff as mature as the Python ecosystem, or more maybe?

[0:32:12.4] AT: That's a good question. I don't think I actually know.

[0:32:16.7] SC: Okay. From a company perspective, where are you in the life cycle of bringing this product to market?

[0:32:22.8] AT: We are very close to general availability. In fact, you can go to our website right now www.light.house and enter your e-mail. We will add you to our special offer list and if you're lucky, you might get one. If not, we will be available for anybody to buy in the not too distant future. We are quite close now.

[0:32:46.8] SC: Nice. I've seen pictures of the device. It looks like – it's not a mobile device. It's stationary. You put it on a counter top or something like that. You either have to be very, very strategic about where you put this thing, or you have envision a world where you've got 10 of these all over the place, like Alexa is becoming. You have one in every room or something like that. Is that the way you're thinking about the world, and you've got – you eventually have a full 3-D and three-color map of everyone's home? Or is it something different?

[0:33:18.9] AT: Actually not. Maybe I'm not doing my job as like CEO of this company and like, "You should have one in every room or something." I actually don't see it that way. I think several in a normal-sized middle-class American house, two to three is probably the right number. You get a ton of value out of one. You get one and you play with it and you're like, "Oh, my God. This is amazing." You can get two and three.

I actually don't think it makes that much sense to have every single room covered. It's usually particular areas of interest. We often see the first one goes in an area that is near the front door. You see like what traffic is coming and going, but you also see a reasonable amount of the floor plan of the home, so you get a sense of what's going on there.

Other common places for Lighthouse to end up in, or in the garage, because often the door might get left open and you want to know if somebody's in there. Or you might have tools out there and children and you want to know, are the kids going out there when I'm not at home, things like that. Or upstairs, in the kids room, or just outside of the kids room to see if they're getting up out of bed in the middle of the night. You would literally just say, "Hey, Lighthouse. Tell me if you see the kids out in the hallway between 10 PM and 6 AM," and it just works.

We need to implement the "Call me" if you see this future. It's coming down the rib. It's those areas of particular interests and it depends on the particular home owner. Another common place is in the living room looking into the area where the dog hangs out, so that you can just get the warm and fuzzy feelings of like, "Hey, what's my dog been up to since I left home?"

[0:34:58.9] SC: If you have this and you point it at the front door, can it effectively track the state of the home and be a general purpose presence detector, like can you track if someone walked in, that person walked out so they're no longer inside at any given time like query it and determine who's in the house?

[0:35:23.6] AT: We can make that query and that does work, but we don't do it with computer vision actually. The reason is there is often many entrances and exits to a home and we don't expect that you buy one Lighthouse for every entrance exit necessarily. The way we do presence-absence detection is with phone presence and absence. GPS is part of that, but also looking at Bluetooth signals coming out of the devices.

[0:35:49.2] SC: I was just going to say I just started playing with the Samsung smart things and it does at the same way and it sucks. It's very coarse.

[0:35:57.7] AT: You have to work hard at it to work well. I mean, there is a big advantage and that we have a Bluetooth signal coming out of the device.

[0:36:04.3] SC: Okay. That's going to work a lot better than the GPS you're within a mile of your house. Therefore, you're in your house.

[0:36:10.4] AT: If you just use location services as provided by the standard phone APIs on its own, it would be hard to make it really good. Now to be fair, also we will not cover the case where you walk out of your house and you go to neighbor's house. It' going to be hard for us to

tell. It will still think you're home, but when you get far enough away then this gets you almost all of the way.

[0:36:34.1] SC: Assuming that you've got your phone with you.

[0:36:36.0] AT: That is correct. Yeah. I mean, that's one of the reasons that the children classifier is a big deal, because they often don't have phones at all. You want to know what they are up to and did they get home on their schedule and so on.

[0:36:50.8] SC: Does Lighthouse have an API? Is it something that you envision people getting and hacking on? Or is it more just the stated use cases SOR?

[0:37:04.5] AT: We have seen a tremendous enthusiasm for adding Lighthouse capabilities to other parts of the IOT world of the smart home. Actuated this smart home device when you see something, or rather that kind of thing. I'm really excited to get to the point where we can actually start to tap into that. We're not there just yet, but it's certainly on the roadmap. We will be deploying something like that, some integration with other smart home capabilities that early adaptors can plug together. We will be providing that sometime this year. It will not be immediate though.

[0:37:44.3] SC: Yeah. What's the long-term view for the company? What are you trying to accomplish?

[0:37:49.0] AT: When I take a step back and look at why Lighthouse exists, the home is a piece of it for sure and it's a very exciting piece, but it's not the only thing. The reason Lighthouse exists is to improve human life by augmenting our physical spaces with useful land accessible intelligence. That's stated very broadly quite deliberately. There is sensors beyond cameras, beyond time of flight cameras and beyond vision generally that are very interesting and that we absolutely should integrate into this thing.

It also goes beyond the home. There's many different AI service domains that are quite interesting to us. We're not spending a lot of time there right there right now, because it's hard enough to do one of these things. We're very, very focused on delivering the home product into the world and having that be a big success and make people's lives better. Once that is

established and growing more less on its own, then it will be time to take our attention to another AI service domain.

[0:38:54.7] SC: What's an example of another one beyond the home that's interesting?

[0:38:58.2] AT: Elderly care is a big deal. It is a particularly big deal and we are particularly well-suited to solve problems in that area. Actually we're starting to see hints of this already, even in the home for aging in place in particular where you have an elderly loved one who maybe they might need to go into a facility, like a nursing care facility, but you want to extend their time in their own home as long as you possibly can.

A system like Lighthouse is actually really good for this use case, because they get a great security camera out of it, or a camera they can see what their dog is up to, or whatever it might be. Then the adult's child gets the early warning system, where you don't have to be looking at it every day. You just say, "Hey, Lighthouse. If you don't see anyone in the kitchen by 8 AM every day, just let me know." Then it might be that they just slept in, but maybe today is a good day for you to call and just see how you're doing.

[0:40:02.5] SC: It seems like there are tons of folks nibbling away at pieces of this space, like how many devices does Amazon have alone? They've got the key thing, which has a camera. They've got the look thing, which is your fashion visual system. They seem to be very gung ho of getting cameras in your home. How does a consumer react to all these people trying to push cameras into their houses and point clouds and all of these stuff?

[0:40:34.0] AT: No, it's interesting. There is a very fundamentally different perspective on the space when you're at a place like Amazon. Their goal is to magnify their marketplace, like they're trying to sell things. That's why they're trying to put a camera into your house so that, "Oh, we can deliver more things to you." Or we can understand that, "Oh, this scarf would look really good on you. I'll try to sell this to you," or whatever it might be. That's legitimately the stated purpose of that device.

With Lighthouse, it's very different. We exist to provide this delightful AI service to you in return for money and that's the end of the transaction. We're not looking to sell you a better hat or something. Taking a step back from all that, it is super interesting what's happening in the home

generally. At the CES in particular, there is this – I almost want to describe it as an epic slugfest between Alexa and Google Assistant to like, "My God. This AI is coming to the home in this particular form." It's really interesting and who knows where it's going to be a year from now. What is very clear is that adding perception capabilities and having the next thing conversational capability is super exciting and that's right where Lighthouse is.

[0:41:57.6] SC: Yeah, there's often this question about is thing X, is it like a product or a feature? What you're doing in a lot of ways is bringing together the vision piece, which is – I guess, I'm wondering long-term like this – did something like Lighthouse and Alexa, do they converge? Do you want to – if Alexa was more open, do you want to have to deal with the NLP? Or do you want the vision to tack onto that, or take advantage of the broader ecosystem? I guess, I'm mostly thinking about this from the perspective of a consumer, like how many of these devices do I want in my house listening to everything?

Already I've got the Google Home and you have to – it has its wake word and Alexa I've got two different wake words in the house. It's like already getting a big maddening.

[0:42:55.7] AT: Yeah. One thing is with Lighthouse, you actually don't talk to the device itself.

[0:43:02.0] SC: Really?

[0:43:02.5] AT: Yeah. Because usually the responses we're providing are video. There is no screen –

[0:43:08.3] SC: You're not talking to your device or something.

[0:43:10.6] AT: Yeah. Usually you're out and about.

[0:43:13.0] SC: That makes a lot more sense.

[0:43:13.9] AT: Usually you're at work, or you're on the train or something and you're just like, "Hey, what did the dog do since I left? Hey, what did the kids do while I was out yesterday?" Then you see the results there. Because it's all about delivering video, answers in video form.

We won't be adding to that confusion about so many different things that can respond to you in the home. I don't have an answer to that problem, but I don't know, you can chat with the Alexa folks, I guess.

[0:43:40.2] SC: Interesting. Interesting. All right, well Alex thank you so much for taking the time to chat with me. I enjoyed learning about you, your background in Lighthouse. Sounds like an interesting space and good luck to you at CES.

[0:43:50.6] AT: Cool. Well, thank you so much. This has been fun.

[0:43:52.4] SC: All right. Thank you.

[END OF INTERVIEW]

[0:43:57.3] SC: All right everyone, that's our show for today. Thanks so much for listening and for your continued feedback and support.

Remember, for your chance to win in our AI at home giveaway head on over to twimlai.com/myaicontest for complete details.

For more information on Alex, Lighthouse or any of the topics covered in this episode, head on over to twimlai.com/talk/103.

Thanks once again to Intel AI for their sponsorship of this series. To learn more about their partnership with Ferrari North America Challenge and the other things they've been up to, visit ai.intel.com.

Of course, we'd be delighted to hear from you either via a comment on the show notes page or via Twitter directly to me at @samcharrington, or to the show at @twimlai.

Thanks once again for listening and catch you next time.

[END]

