**FEI-FEI LI**: (VOICEOVER) Whenever humanity creates a technology as powerful and potentially useful as AI, we owe it to ourselves and our future generation to make it right.

**[MUSIC]**

**KEVIN SCOTT:** Hi, everyone. Welcome to *Behind the Tech*. I'm your host, Kevin Scott, Chief Technology Officer for Microsoft.

In this podcast, we're going to get behind the tech. We'll talk with some of the people who have made our modern tech world possible and understand what motivated them to create what they did. So join me to maybe learn a little bit about the history of computing and get a few behind-the-scenes insights into what's happening today. Stick around.

**[MUSIC]**

**CHRISTINA WARREN:** Hello and welcome to Behind the Tech. I'm Christina Warren, Senior Cloud Advocate at Microsoft.

**KEVIN SCOTT:** And I’m Kevin Scott.

**CHRISTINA WARREN:** And today, we're continuing our conversation about AI, this time with Stanford professor and researcher, Fei-Fei Li.

Now, if you didn't catch the last episode, be sure to check it out, because Kevin has a conversation about AI with another Stanford academic and researcher, Surya Ganguli. And you guys went deep.

**KEVIN SCOTT:** Yeah, it's sometimes very hard to contain myself. We got super technical just because it was super fascinating what we were chatting about.

**CHRISTINA WARREN:** Yeah, and so, one of the things that we talked about last time, Kevin, was the need for inspiring and positive stories about AI. And I think that you and I, I think we made a promise to write a screenplay.

**KEVIN SCOTT:** Yeah, in our copious free time, right? (Laughter.)

**CHRISTINA WARREN:** Exactly. We have so much of it. But, you know, I have been thinking about the screenplay. And so, recently I was reading an article about AI and animals and how they think that they're only about a decade out from having a kind of language translator.

**KEVIN SCOTT:** So, you mean like the dogs from the Pixar film Up? Squirrel?

**CHRISTINA WARREN:** Yes, yes, yes, the small mailman smells like chocolate. (Laughter.) So, the research might have applications to animal welfare where, for instance, they might be able to use AI to track the faces of sheep or cows in order to detect pain and illness, and then provide faster medical treatments.

So, they've also done some really incredible studies with prairie dogs, who they believe have a very complex language system.

And so, that would be really tempting, in my opinion, to kind of write a show with the typical trope of the talking family pet, you know, like Mr. Ed. But, you know, Grumpy Cat, may she rest in peace, you know, maybe she wasn't grumpy at all, maybe she was like happy-go-lucky.

But kind of back to the prairie dogs, you know, I do think that that -- there's -- that's one that hasn't been done yet. You know, we could maybe do like a *Little House on the Prairie* remake but with AI.

**KEVIN SCOTT:** Yeah, you know, I we're trying to make a stab at comedy here, but funny enough, when Fei-Fei and I were chatting before the podcast, she was telling me about this Chinese company that does face recognition for cows. So, like you may not be that far off. But despite that, like maybe we should stick to our day jobs.

**CHRISTINA WARREN:** Maybe, I guess. Okay, all right, so we should definitely meet Fei-Fei Li.

**KEVIN SCOTT:** Yeah, let's do it.

**[MUSIC]**

**KEVIN SCOTT:** Next up, we'll meet with Fei-Fei Li. Fei-Fei is considered one of the pioneering researchers in the field of artificial intelligence. She's a computer science professor at Stanford University and the co-director of the Human-Centered AI Institute there.

Fei-Fei has served as the director of Stanford's AI lab, and during a recent sabbatical, she was a VP at Google, serving as Chief Scientist of AI and Machine Learning at Google Cloud.

**[MUSIC]**

**KEVIN SCOTT:** So, thank you so much Fei-Fei for coming in today. I've been wanting to chat with you for a really long time now.

**FEI-FEI LI:** Likewise, Kevin. Thank you for inviting me.

**KEVIN SCOTT:** So, I usually start these things by trying to understand a little bit of the story of the folks that we're chatting with. And I'd be really interested to understand like how you started to really get seriously interested in computer science.

**FEI-FEI LI:** So, I came to computer science through a pretty convoluted detour. So, I was always kind of a STEM kid, so to say. So, I was interested in the nature, in the stars, the animals and all that. But my first passion, first love was physics.

**KEVIN SCOTT:** Awesome.

**FEI-FEI LI:** So, starting in junior high and then high school, I was just passionate about physics, studying relativity, you know, reading --

**KEVIN SCOTT:** And what was it about -- was it that you -- like, it gave you a lens to understand the world? --

**FEI-FEI LI:** Yeah.

**KEVIN SCOTT:** --Was it that you just liked the mathematics of it? What was the thing?

**FEI-FEI LI:** I think it's the combination of the imagination and the mathematical rigor, but it's really about peeling off question after question to go after the very original questions.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** Right? Like, where do I come from? Where do humans come from? What's human made of? Where do atoms come from? Where do the first atoms come from? You go to Big Bang.

**KEVIN SCOTT:** So, basically, you're infinitely curious?

**FEI-FEI LI:** Yeah, so I was. So, physics was my love, and I majored in physics at Princeton when I went to college.

**KEVIN SCOTT:** Awesome.

**FEI-FEI LI:** And, you know, Princeton is the mecca for physics. And first day in freshman year physics class, the professor said, "This is the very lecture hall that Einstein was sitting in." (Laughter.) It was just like a dream, right?

So, but the -- around sophomore, junior year in college, I started reading books about these great, great physicists of 20th century like Schrödinger and Einstein. And as I read, I noticed that towards the later half of their life, their interest turned from the physical atomic world to the more life science world.

They're starting to ask the questions of origin of life, intelligence. And that really piqued my interest. I started to get very interested in the question about intelligence. So I joined neuroscience research. And I was literally as a summer intern research student, literally recording from mammalian brain and listening to the neurons, seeing the world.

**KEVIN SCOTT:** So fun.

**FEI-FEI LI:** Yeah, so, I decided to apply for grad school, and even there, I chose to go to Caltech because I was able to find two advisors, one in neuroscience and one in -- now we call it AI, but at that time, we call computer vision.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** --to do my PhD study in that combination.

**KEVIN SCOTT:** And why vision?

**FEI-FEI LI:** Yeah, good question. (Laughter.) So, well, like I said, my first experiment in neuroscience was recording from cats' neurons -- and watching their neural activity when cats see the world, you know, see oriented edges, see complex features.

But, really, if you think about vision, I want to tell you a story. 540 million years ago, the world is very different. It was mostly water and simple animals floating on earth. And there were only a few species.

And life was chill, you know, you just hang out by floating. (Laughter.) And then what zoologists have found, or evolutionary biologists have found this, this incredibly mysterious phenomenon called "Cambrian explosion," within a short 10-million-year span, in the history of earth or in the evolution sense, it's such a narrow slice of time.

The animal kingdom just exploded. Many, many animal species are created or evolved. And people called this the "Big Bang of evolution," and no one understood why.

So, why, from 540 million years to 530 million years ago this big bang of evolution happened?

So, fast forward. A decade ago, a young evolutionary biologist called Andrew Parker from Australia studied a lot of fossils and conjectured to become a very convincing theory that it was the onset of vision.

**KEVIN SCOTT:** Really?

**FEI-FEI LI:** It was the first animal, some kind of floating trilobite developed a pin-hole structure. It's very simple, it literally just collected some light.

Once you see light, life changes. You become active. You see your food. And you can hide and escape to become someone else's food.

So, it becomes an evolutionary arms race for animal species. And with that kind of active lifestyle, so to say, because of vision, animals started to evolve much faster.

**KEVIN SCOTT:** Interesting.

**FEI-FEI LI:** Ans, fast forward 540 million years later, visual intelligence is the most fascinating to me and complex sensory system of human brain. Half of our brain are involved in visual processing and understanding. That's a lot.

**KEVIN SCOTT:** And it's super interesting because I think a lot of people, you know, if you're trying to just sort of reflect on your own intelligence, like a lot of that is about language, like in fact like even the – you know sort of the meta process of thinking about your intelligence is linguistic, like you're sort of having this dialogue with yourself and with everyone, but your suggestion is that like vision is like the more, you know, sort of fundamental thing, maybe.

**FEI-FEI LI:** So, I'm not saying that language is not important and it's a unique part of human intelligence, but I recommend you read a book by Alison Gopnik called *Scientists in a Crib*, and she's a developmental psychologist and a philosopher who studies babies, very young babies.

So, when you say language, I just want to challenge you. Are babies intelligent? Before they developed language, right? In fact, they're the most fascinating creatures we create because in our first two years of life, which language is not the primary tool, they are just curious creatures exploring, understanding and interacting with the world. They develop the theory of the other mind.

They develop the sense of objects. They develop social intelligence. They do face recognition. They navigate. They manipulate, they crawl, they understand space.

This is all without language. So, I just want to highlight how incredibly deep, important and useful visual intelligence is. And, of course, as soon as language gets developed, you can see the interaction between vision and language, one of the most exciting areas of research that I'm doing right now is the interplay between vision and language. The vision is, for me, is just highly fascinating.

**KEVIN SCOTT:** So, tell us a little bit about your PhD work. So, like, you're in this program, you've got a neuroscience advisor and like basically a computer vision advisor. And so, like, what does your dissertation research look like?

**FEI-FEI LI:** Yeah, (Laughter.) so that was a great question. So, I literally did a combination of cognitive neuroscience and computer vision.

So, on the cognitive neuroscience part, I started my PhD in the first year of 21st century, 2000. And little did the public know that even today's AI revolution owes a lot to the incredible advances in cognitive science, starting from '70s, '80s, and going well into the '90s, because we're mapping out some of the incredible capabilities of human intelligence system, including vision.

And one of the most fascinating area of vision at that time I was studying is our ability to understand natural world. We -- an earlier study by a English scientist living in France, Simon Thorpe, shows within 150 milliseconds of seeing a complex visual, seeing humans are already capable of understanding if this thing contains an animal with it-- or it doesn't. And here, we're talking about all kind of potential animals in all kind of environment.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** So, that processing was fascinating. So, one of my study in PhD time was actually to quantify how much we see at the moment we open our eye from objects to people to movements --

**KEVIN SCOTT:** And how do you do that experiment?

**FEI-FEI LI:** Oh, that's fun. So, in cognitive neuroscience, it's called psychophysics. So, what I would do at that time is collect hundreds of photos from Flickr, actually, and these photos are all daily user uploaded photos. So, it goes from like birthday parties to, you know, surfing in the ocean, to all kind of topics.

And then we put a program where you code up a program, and then you put a human subject in front of a computer screen and then you flash the photo quickly and you control the amount of time you flash the photo.

We literally went down to 27 milliseconds, all the way to 500 milliseconds to show the photo to the human. And then we control how long the picture gets seen and then we ask the people to write down what they see.

**KEVIN SCOTT:** Interesting.

**FEI-FEI LI:** And we pay $10 for the undergrads who participate (laughter) in the experiment. And then we collect a lot of data, of course there is scientific rigor had to kick in.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** And then we statistically analyzed this and understood it was one of the first studies that ever quantified how much people see within literally a glance of a thing.

**KEVIN SCOTT:** That's really interesting. So, like, 150 milliseconds, like, your brain can see this image and decode enough of it where you can sort of explain back --

**FEI-FEI LI:** Yeah.

**KEVIN SCOTT:** --what's in it?

**FEI-FEI LI:** Exactly. Test it yourself, just open your eye and close it. That will be longer than 150 milliseconds, but within that very short amount of time, the comprehension you have of the visual world is so rich. And that inspired my AI work because at that time, most of computer vision was still you know recognizing letters or --

**KEVIN SCOTT:** Oh, I remember.

**FEI-FEI LI:** Right? And -- or --

**KEVIN SCOTT:** And they were non-neural-network models and --

**FEI-FEI LI:** Right. Well, neural network was invented, in fact, my first course at Caltech was called Neural Networks. But it was working on very simple stimuli like digits and numbers. And then computer vision was still trying to understand edges.

And then my advisor, Pietro Perona, was one of the pioneering scientists in computer vision who said, "Why don't we venture into real-world object recognition?' And with my study on the neuroscience side, we also have evidence that this is what humans are capable of and are good at and we really have to move computer vision towards that human capability.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** So, I started my AI research in enabling computers to see and understand everyday objects.

**KEVIN SCOTT:** But that's a big leap. I mean, we sort of take for granted that image classifiers are reasonably good now, but --

**FEI-FEI LI:** Yeah.

**KEVIN SCOTT:** -- like, when you're doing this back in 2000 to like have the idea that I want to go from this, like, relatively sort of simplistic and not to, like, denigrate in any way all the computer vision research, because like some of the best people I know -- like the smartest people I know were doing that work. But it's a big leap from that to like I want to do like whole object recognition inside of images.

**FEI-FEI LI:** It was a leap. I mean, I didn't single-handedly do it. Like I said, there were a few incredibly forward-thinking scientists, Jitendra Malik, David Lowe, Pietro Perona, they're starting to thinking that way.

But we've as a field made mistakes and had detours, right, at that time. We were thinking about how to mathematically construct those handmade models to describe objects, and that took years and years to do. And it didn't deliver the results we want.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** And one of the projects, it's really ironic, but also fun to reflect back is that data was so scarce at that time that my very first object recognition project is called One-Shot Learning is to work on a setting that we only have one or two picture to train the algorithm.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** And today, think about the big data age. And what I did for ImageNet, it was almost the polar opposite.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** But it's a capability humans have.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** And we try to replicate that.

**KEVIN SCOTT:** Yeah, and I want to come back to that in a minute, but like you brought up ImageNet, so like this is one of the things you're most well known for.

And, you know, as I, listen to you talking about your PhD work, it sort of seems like a natural extension to an extent of what you're doing. So, like, for the audience, like, why don't you describe what ImageNet is?

**FEI-FEI LI:** Okay. So, ImageNet was a project we started in 2007 and more or less completed in 2009. The end result is at that time the largest database of natural object images in the world.

It consisted of 15 million images organized in everyday English language of 22,000 vocabularies, mostly nouns. And we collected this data set for about three years by labeling, cleaning, sorting almost a billion Internet pictures.

And what ImageNet did is it provided one of the most critical ingredients as data for enabling neural network architecture to train, and that was the onset of the deep learning revolution.

**KEVIN SCOTT:** Yeah, I mean, like, you sort of are being a little bit understated about it. But, like, it's almost impossible to imagine, like, how that rapid iteration loop with exploration of these DNN architectures and, like, the techniques that we develop to train them quickly on GPUs and whatnot, like, none of that really could have happened without this big database of training data.

**FEI-FEI LI:** Well, thank you, that's very nice of you for saying that. And we designed and developed ImageNet because we believed strongly in around 2006, 2007, that we have to hit the reset button for machine learning. That the work we have been doing the past few years exploring different models didn't quite work for the scale and the scope of our real natural visual world of so many different objects and the varieties they represent.

So, my students and I conjectured that the way to really think about modeling objects through machine learning techniques is to think through data.

That was a pretty bold statement at that time because, at that time people were constructing small probabilistic models through hand design and a lightweight training of parameters. So, for us to go in to say that, hold on, let's just rethink about this whole thing through a data point of view was kind of a, you know, a minority way of thinking. (Laughter.)

**KEVIN SCOTT:** Yeah, but you know again, like just sort of absolutely necessary. And like, curious what you think about -- we're sort of in this state right now where we're beginning to see people do really interesting things with reinforcement learning and unsupervised learning where you're getting a little bit away from requiring so much explicitly labeled data.

So, like, the data is still very important, but like you don't have to go through this exercise of annotating, like, oh, there's a cat in this image, there's a red ball in this image.

**FEI-FEI LI:** Yes, right.

**KEVIN SCOTT:** And like the results actually are in -- the interesting ones that have come out over the past couple of years have mostly been natural language things with these big, unsupervised models. Like, do you think that's sort of a trend that will continue across a bunch of different domains?

**FEI-FEI LI:** Absolutely. I think this is very exciting. I think, you know, if you reflect on human intelligence, our way of learning is very multi-dimensional, right? We do have training and supervision-based learning, especially when a teach a kid manner, it seems lots of supervision is needed. (Laughter.) But children and grown-ups learn from trial and error, learn from few-shots, learn in unsupervised setting, learn with rewards and punishments sometimes.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** So, that kind of flexibility is clearly critical, and evolution has built into human intelligence. And for machine intelligence to become more robust to serve human world better in different settings, I think that kind of unsupervised learning or few-shot learning or reinforcement learning is absolutely needed.

I'll give you an example, that I work with healthcare industry a lot because part of my research these days deals with AI in healthcare. And one of the settings that we work with is senior wellbeing and senior safety. And fall detection is a huge thing for seniors.

In fact, falling accounts for billions and billions of dollars of medical spending for American seniors and you know, it can be fatal and even if not fatal, can cause a lot of pain and issues for our aging population.

Well, when it comes to falling, it's a rare event. We immediately run into a lack of data problem when we are working with the doctors, right? It's very, very hard to collect an ImageNet of seniors falling, and you don't want to.

**KEVIN SCOTT:** Right.

**KEVIN SCOTT:** And, like, the seniors don't have cameras in their homes and -- yeah.

**FEI-FEI LI:** Right. And they fall in such different ways and the situation is complex. So, when you want to work in these critical issues, you're immediately in a so-called few-shot learning situation. And you probably have to consider transfer learning, consider, you know, simulated learning and all that.

So, it just shows that the field now -- I'm very glad to see is moving beyond just large data supervised learning.

**KEVIN SCOTT:** You mentioned healthcare, which is, I think like one of the most promising focus areas for AI right now. And you know, I know that you're one of the co-directors of the Stanford Human-Centered AI Institute and --

**FEI-FEI LI:** And you're one of our advisors. (Laughter.)

**KEVIN SCOTT:** So, how are you thinking about what we need to do to get AI to better serve the interests of everyone?

**FEI-FEI LI:** Well, Kevin, that's a big topic. And I think that's a really important topic, right? Whenever humanity creates a technology as powerful and potentially useful as AI, we owe it to ourselves and our future generation to make it right.

So, first of all, I think the institute that both of us are involved in is really laying out a framework of thinking about this, and the framework is human centered. It's that from the get-go, from the design and the basic science development of this technology all the way to the application and impact of this technology, we want to make it human benevolent.

And with this framework in mind, we have at Stanford, this institute works on three principles -- founding principles to cover different aspect of human-centered AI.

The first principle is actually what we've been talking about is to continue to develop AI technology, basic science technology, that is human inspired and betting on the combination of cognitive science, psychology, behavior science, neuroscience, to push AI forward so that the technology we will be using have better coherency or better capability to serve human society.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** So, that's the first principle. Second principle is I would love to hear your thoughts. You know, you and I are trained as a generation of technologists that the technology is solidly considered a engineering field or computer science field. But I think AI really has turned a chapter. AI is no longer a computer science field.

**KEVIN SCOTT:** Correct.

**FEI-FEI LI:** AI is so interdisciplinary today. In fact, some of the most interesting fields that AI should really contribute and also welcome to join force are social sciences and humanities. And at Stanford, we're already seeing the collaboration between AI researchers with economists, with ethicists, philosophers, education experts, legal scholars and all that.

To do this, our goal is to understand what this technology is really about, understand its impact, but also forecast, anticipate the perils, anticipate the pitfalls, anticipate unintended consequences, and really with the eventual goal of guided and recommend policies that are good for all of us. So, that's the second principle is really understand, anticipate, and guide AI's human and societal impact.

The third and the last but not the least principle is something I know you and I feel passionate about is really to emphasize the word "enhance" instead of "replace."

**KEVIN SCOTT:** Correct.

**FEI-FEI LI:** Because AI technology is talked about as a technology to replace humans. I think we should stay vigilant about job displacement and labor market, but the real potential is using this technology to enhance and augment human capability, to improve productivity, to increase safety, and to really eventually to improve wellbeing --

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** --of humans. And that's what this technology is about. And here, we're talking about healthcare. Another vertical that we put a lot of passion and resource in is education. Sustainability. Manufacturing and automation, these are really humanly and societally important areas of development.

**KEVIN SCOTT:** Yeah. Well, just sort of sticking with healthcare and like your eldercare example, like, this is something that I don't think a whole lot of people spend time thinking about unless they're taking care of an elderly parent or relative.

Like, we're not thinking about, like, how systemically we can make the lives of elderly people better. And, like, we're certainly not thinking about the big demographic shifts that are about to come --

(Crosstalk.)

**FEI-FEI LI:** Oh, my god, it's going to come globally.

**KEVIN SCOTT:** Yeah, globally. I mean, so, you and I have chatted about this before, but you know, we sort of see in almost all of the industrialized economies, but also in Japan, Korea, and China --

**FEI-FEI LI:** Yeah, absolutely.

**KEVIN SCOTT:** --you have this very large bubble of working-age population that's getting older and older. And we just don't have high enough fertility rates in these younger generations to replace it.

So, at some point, like, we -- across the entire world, we're going to have far more old people than we will have working-age people. And you have like a couple of big questions when that happens. Like, who takes care of all the old people and like who's going to do all the work? And, it's actually not far enough away that we can not think about it.

**FEI-FEI LI:** 2035 is --

**KEVIN SCOTT:** Yes.

**FEI-FEI LI:** -- I think the -- we have to find the actual number, but the last Baby Boomers become the aging population, the youngest. So, we -- we're very close to that. And also, to do this research in aging population, I spend a lot of time in senior homes and senior centers. One thing I learned as a technologist is that we should really develop the kind of empathy and understanding of what we really are working on and working for. For example, I cannot tell you how many Silicon Valley startups are there to create robots as senior companions. And when some of them feel robots can replace family, nurses, friends, I really worry. And I really want to encourage these entrepreneurs to spend a lot of time with the seniors.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** One thing I learned a lot about wellbeing with aging population is dignity, social connection is part of the biggest part of aging. And so my dream technology is something that you don't notice, but it's quietly there.

**KEVIN SCOTT:** Correct.

**FEI-FEI LI:** To help, to assist, to connect people, to ensure safety. Rather than this big robot, you know, sitting in the middle of the living room and replacing the human connectivity.

**KEVIN SCOTT:** Yeah, it's really funny that you're bringing all of this up. I'm writing a book right now on, why I think people should be hopeful about the potential of AI, like, particularly in rural and middle America. And for the book, I went back to where I grew up in rural central Virginia in, like, this, you know, very, small town.

And I visited the nursing home where three of my grandparents spent the last chunk of their life. And I was just chatting with some of the people there.

And I asked them, you know, the nurses and the managers in this place, like, you know, what -- do you think AI -- and like, when I say AI, like, the vision that conjures is, like, oh, there's going to be some human equivalent android coming in. And they'd be like, no, the residents would be terrified -- by this thing. Whereas like they've got a bunch of thing -- like, dispensing medicine, for instance.

**FEI-FEI LI:** Exactly.

**KEVIN SCOTT:** Like, you know, when you're elderly, like, you're taking this, like, complicated cocktail of medicines and, like, getting it dispensed in the right amounts at the right time through the day, making sure that you actually take the medicine.

Like, that's a problem that we could solve with AI-like technologies, like, you know, combination of robotics and computer vision. But it wouldn't be like this talking, walking, you know, robot. It would be, like, a set of things that sort of disappear into the background and just sort of become part of the operation of the place.

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** And, like, that I think we should have more ambition for that sort of thing rather than this --

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** You know?

**FEI-FEI LI:** That's why Stanford HAI wants to encourage that. The best technology is you don't notice the technology.

**KEVIN SCOTT:** Yes.

**FEI-FEI LI:** But your life is better.

**KEVIN SCOTT:** Yes.

**FEI-FEI LI:** That's the best technology.

**KEVIN SCOTT:** I could not agree more.

**FEI-FEI LI:** And also just talking about the rural America, this is something I feel passionate about, and I have a story to share with you. So, you probably know that I co-founded and chair this nonprofit educational organization called AI4ALL, right?

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** It started as a summer camp at Stanford about five years ago to encourage diversity students to get involved in AI, especially through human-centered AI, studying and research experience to encourage them to stay in the field and then our goal is in ten years, we would change the workforce composition.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** Now, it became a national nonprofit and seed granted by Melinda Gates and Jen-Hsun Huang Foundation and --

**KEVIN SCOTT:** That's awesome, I didn't know Jen-Hsun was involved. That's great.

**FEI-FEI LI:** Yeah. It's Jen-Hsun & Lori Huang Foundation. And this year, we're on 11 campuses nationwide. One of the populations we put a lot of focus on in addition to gender, race, income, is geographic diversity and serving rural community, for example, our CMU campus is serving rural community in Pennsylvania. We also have Arizona campus.

One story actually came out of our Stanford camp is Stephanie. Stephanie is still a high school junior now. And she grew up in the backdrop of strawberry field in rural California.

In a trailer park with a Mexican mom. And she come from that extremely rural community, but she's such a talented student and has this knack and interest for computer science.

And she came to our AI4ALL program at Stanford two years ago. And after learning some basics about AI, one thing she really was inspired is she realized this technology is not a cold-blooded, just bunch of code. It really can help people. So, she went back to her rural community and start thinking about what she can do using AI to help. And one of the things she came up with is water quality.

**KEVIN SCOTT:** Yes.

**FEI-FEI LI:** Really matters to her community. And so, she started to use machine learning techniques to look at water quality through water samples.

And that's just such a beautiful example. I just love her story to show that when we democratize this technology to the communities, the diverse communities, especially these communities that technology hasn't reached enough in, the young people, the leaders and the citizens of this community will come up with such an innovative and relevant ideas and solutions to help those communities.

**KEVIN SCOTT:** Yeah. And I think that getting this technology democratized is sort of a one-two punch.

So, like, there's the tactical things that you have to do. So, and open source and, like, making sure that the research is open and freely available and being able to run these things on cloud platforms, you know, and like all of that's super important. It's actually amazing, like, how --

**FEI-FEI LI:** Cloud and Edge.

**KEVIN SCOTT:** Yes, cloud and Edge for sure. And you know, it's really amazing like how much is possible now. Like, I know you probably have this all the time, it's like you're sitting in 2019 and you know, seeing what your students can do. And you sort of compare that to what you could do in 2000 it's, like, you know?

And, like, that's because you have bright students, but it's also because, like, the tools that they're using are, like, incredibly sophisticated now.

But that's only half of the story. Like, the other half, and like I'm so glad that that you're doing this nonprofit work because if we really want the benefits of this technology to be, you know, sort of equitably and widely distributed, you have to have people who have a connection to the communities and the human beings that the technology needs to serve.

**FEI-FEI LI:** Absolutely. Absolutely.

**KEVIN SCOTT:** Because -- and it's not that anybody's bad, it's just like if you don't have that context and that empathy, like, you just don't really know what to do or maybe even how to do it.

**FEI-FEI LI:** Absolutely. We had an alum who one of the grandparents unfortunately passed away due to a delay of ambulance service. Now, she's working on machine learning and optimizing ambulance dispatch.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** I think that's why we need all walks of life is they because they bring the understanding and empathy you said and also the experience to innovate and create in ways that just one slice of people couldn't possibly cover.

And you said the right thing, people is at the heart of all this. When AI4ALL was founded, our slogan was, "AI would change the world, who will change AI?"

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** That is the core of this problem.

**KEVIN SCOTT:** Yeah, it's awesome. So, what are you most excited about? And I'll ask it two different ways. So, like, what are you most excited about from a research perspective right now in AI? And, like, what are you most excited about from a social good perspective, like --

**FEI-FEI LI:** And, hopefully, they actually are not mutually exclusive.

**KEVIN SCOTT:** Yes, I very much hope that they're not. (Laughter.)

**FEI-FEI LI:** I think from basic research science point of view, there is one direction that I'm exploring with my collaborators at -- students at Stanford that really excites me. And it goes back to what we were saying about the babies and *Scientists in the Crib* because early childhood is this rich period of learning the world that is in such fascinating ways.

This is where you're not labeling a thousand cat images and showing it to a baby and saying, "Cat, cat, cat," right? That just doesn't work. They're just exploring out of curiosity and all that.

So, there's a project at Stanford I'm involved in and I have students working on it, it's curiosity-based learning is where we design machine learning agents and put them unfamiliar environment and they have capability to interact with the objects in the environment and watch how the agent, through these kind of curiosity-based learning, develop capabilities of recognizing objects or understanding physical properties of objects.

**KEVIN SCOTT:** And is this a variation of reinforcement learning where --

**FEI-FEI LI:** It uses reinforcement learning. It uses, definitely, deep learning as the early representation of the world. Deep learning is very useful. It's a combination of deep learning and reinforcement learning. It's curiosity driven, so --

**KEVIN SCOTT:** And how do you articulate the -- so, like, there must be some metric for it, right?

**FEI-FEI LI:** Right. Curiosity is expressed through your known model of the world and the difference you observe.

**KEVIN SCOTT:** Oh, interesting.

**FEI-FEI LI:** And for babies, it's the same, right? If they keep seeing the same thing, they get bored. So, they want to explore different aspects. So, they want to create new things.

So, if you give a baby a ball, maybe he or she would first look at the ball and then get a -- he or she would drop the ball. If you give him or her two balls, they will start banging the two balls. So, these are the different aspect of --

**KEVIN SCOTT:** Oh, interesting.

**FEI-FEI LI:** -- of interacting with the world. So, we start seeing that. And it's still early research, but what I would love to see is behavioral patterns emerge from the machine learning agent, and then we can do human experiment to contrast and compare and see if we can improve our machine learning algorithm, but also to see what emerges for machines that are different from humans.

**KEVIN SCOTT:** And do you imagine in this research that the models will be very large, just because you want something that's sort of expansive and has the room to learn different sorts of representations in this space? Or --

**FEI-FEI LI:** So, this is where it's very different from the brain. We start small. The models, because they simulated environments, they're pretty simple, so a couple of objects with simple shape and color and material.

But we want to grow the world -- the machine agent world and I'm not going to be surprised if this model becomes larger and larger.

**KEVIN SCOTT:** Right. Yeah, I mean, the reason I ask is like one of the things that really has started to intrigue me over the past few years is -- and like I think this has sort of been true for like, the past decade or so.

The things that have been making the fastest progress in AI are things that have some sort of connection to one or more things that are growing exponentially fast.

**FEI-FEI LI:** Chips.

**KEVIN SCOTT:** So, like compute --

**FEI-FEI LI:** Data.

**KEVIN SCOTT:** Compute and data have been the two big things that are, you know, that are driving not -- they're not the sole things -- they're not driving progress at all, actually, they're facilitating, very rapid progress. And so, like, I'm always looking for that connection when --

**FEI-FEI LI:** Yeah, on the other hand, the human brain operates on less than 20 watts.

**KEVIN SCOTT:** Yeah, I know. It's a brilliantly efficient thing.

**FEI-FEI LI:** And, exactly. It doesn't take that many neurons to get the first, you know, impression of the world when you open your eyes.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** So, there are some really interesting contrasts in biological intelligence and machine intelligence.

**KEVIN SCOTT:** Yeah, I was -- I'm probably getting the details on this wrong, but I remember like even a couple of years ago reading a, like, a little short note in *Science* or *Nature* about how someone had used FMRI to map out the primate neural network –like, biological neural network that does face recognition. And it was like tiny, like, little -- little bitty network --

**FEI-FEI LI:** It's called -- yeah, the central area is called FFA, in fact, that was -- in 1990s, late 1990s, the MIT researcher Nancy Kanwisher and many of her colleagues were at the forefront of that study and really give rise to a lot of neural correlate belief that there are areas of brain with those kind of expertise.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** And they're not that huge.

**KEVIN SCOTT:** Yeah, and so before we get onto the social stuff, which I'm super interested in, like, tell me a little bit more about this work that you're doing that sort of blends vision and language together, because that seems really quite exciting.

**FEI-FEI LI:** Yeah. So, it actually is a continuation or step forward from ImageNet. If you look at what ImageNet is, for every picture, we give one label of an object.

Fine, that's cool. You have 15 million of them, it becomes a large data set to drive object recognition. But it's such an impoverished representation of the visual world.

**KEVIN SCOTT:** Right. Yes.

**FEI-FEI LI:** So, the next step forward is, obviously, to look at multiple objects and you know, be able to recognize more. But what's even more fascinating to me is not the list of 10 or 20 objects in a thing, it's really the story.

And so right after the bunch of work we have done with ImageNet around 2014 when deep learning was, you know, showing its power, my students and I started to work on what we call image storytelling or captioning. And we show you a picture, you say that two people are sitting in a room having a conversation. That's the storytelling. And that is a sentence or two, right?

And, honestly, I'll tell you, Kevin, when I was in grad school in early 2000, I thought I wouldn't see that happen in my lifetime because it's such an unbelievable capability humans have to connect visual intelligence with language --

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** -- with that. But in early 2015, my group and my students and I published the first work that shows computers having the capability of seeing a picture and generate a sentence that describe the thing.

And that's the storytelling work. And we used obviously a lot of deep learning algorithm, especially on the language side, we used recurrent models like LSTM to train the language model, whereas on the image side, we used convolutional neural network representation. But stitching those together and seeing the effect was really quite a “wow-wee” moment.

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** I couldn't not believe that I saw that in my lifetime, that capability.

**KEVIN SCOTT:** Yeah, I mean, I sort of wonder, like, whether or not these big, unsupervised language models right now, these transformer things that people are building, the models that come out of them, like, have such -- they're just very large and, like, there's not much, you sort of barely have, like, any signal in the parameters at all.

It's, like, just diffuse across the entire model -- I just wonder, like, whether getting, like a vision model coordinated with training these things is going to be the way that, like, they more concisely learn.

**FEI-FEI LI:** Oh, I see. Well, yeah, I mean, human intelligence is very multi-modal.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** So, multi-modality is definitely not only complementary, but sometimes is more efficient. We should also just recognize that, by and large, these storytelling models are still fitting patterns. They lack the kind of comprehension and abstraction and deep understanding that humans have.

They can say two people are sitting in a room having a conversation, but they lack, the common-sense knowledge of the social interactions or you know why are we having eye contact or whatever, right? So, there is a lot more deeper things going on that we don't know how to do yet.

**KEVIN SCOTT:** And so on the social side, like, what are you excited about? And obviously, like, you've already talked about a ton of it. Like, you're doing this really interesting work in healthcare, HAI, like, I think both in and of itself and as a, you know, just sort of an example and role model for the, like, wider academic world is, like, a fantastically good thing and sort of calling out that, like, this has to be, you know, inclusive and multi-disciplinary. But, like, what are you hopeful about in the future?

**FEI-FEI LI:** Right. So, oh, god, so many things, right? Even on the healthcare side, we just touched the aging population, but what I really feel so passionate about is my collaborator, Dr. Arnie Milstein and I really see that while there's a lot of talks about AI and healthcare, much of that is in the diagnosis and genomics side, but there's a huge open issue in the care delivery side. In fact, in America, that medical-error-induced fatality is causing a quarter of a million lives every year, right?

**KEVIN SCOTT:** Right.

**FEI-FEI LI:** Hospital-acquired infection alone kills three times more people than car accidents. And you mentioned smart sensors. The same technology we're using for self-driving cars between smart sensors and deep learning algorithms can become so critical and helpful in improving care quality from our surgical rooms, to ICUs, to senior homes.

And so we are passionate about continue working in that space and to change healthcare delivery quality. In addition to that, both by working Stanford's HAI and AI4All, what I really want to do is to create this platform.

I cannot possibly do all the work. I don't possibly have all the good ideas. But creating the platform to welcome the kind of talented people, thinkers, students, and leaders, practitioners, policy makers, civic society, to participate in this effort and movement I think will be critical for our future. You know, we talk about how to predict our future, well, the best way to predict is to create.

**KEVIN SCOTT:** Yeah, oh, god, I couldn't not agree more. And like showing everyone that you know there are far more hopeful paths than there are you know, sort of pessimistic ones, like, gives everyone both inspiration and permission to go off and, like, create that more hopeful future.

**FEI-FEI LI:** Yeah, and I particularly want to encourage and inspire people. You do not have to be a coder to join AI and to change AI.

I think that myth from Silicon Valley that you have to be a, you know, a coder from 11 year old and know TensorFlow or whatever inside out in order to be part of this AI movement, that's absolutely not true. We need artists, we need writers, we need social scientists, we need philosophers.

**KEVIN SCOTT:** I totally agree. We need more people involved in more ways with this technology than we ever have in the, like, sort of lifetime of digital technologies.

And, like, I would even argue that AI itself is making the task of developing things, like the engineering task, different and more inclusive --

**FEI-FEI LI:** Yes.

**KEVIN SCOTT:** Like, you and I, you know, got into, you know, sort of computer science because, like, you know, we have a certain sort of analytical way of seeing the world and, like, we really enjoy, like, all of the apparatus of that analytical world. But, you know, there are these machine teaching systems where, like, you're going to be able to, you know, sort of rather than tell the computer what to do and these, like, minute, you know, step-by-step algorithmic ways, you're going to be able to teach a computer how to do something.

And, like, that is a really -- like, much broader -- mode of, you know, sort of building these bits of technology.

**FEI-FEI LI:** I can't tell you how many artists have reached out to me and to Stanford HAI about AI helping the creative process. They're so excited.

**KEVIN SCOTT:** Yeah, it's super exciting. One of my -- you know, I hope to be able to get him on the podcast at some point, but there's this fantastically talented young jazz musician named Jacob Collier, who, he's like a genius with harmonic theory.

And he does, like, all of these, like, super interesting, innovative arrangements. And, like, he got famous by recording these layered things that he was making on YouTube. But he, like, really enjoys performing these things live. And so he's been collaborating with this -- with this really talented engineer at MIT to build instruments where he can reproduce some of this self-harmonization stuff --

**FEI-FEI LI:** Oh, that's so cool.

**KEVIN SCOTT:** -- and like AI is going to do nothing but help him, like, be able to, like, deliver these richer experiences to his audience. I mean, it's just --

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** -- it's amazing. Like, I'm -- this is the stuff that makes me really -- really super excited. (Laughter.)

So, like, one last question before we wrap up. So, I know you're a mom and you're -- you've got a nonprofit, you're institute director, you're a professor --

**FEI-FEI LI:** Researcher.

**KEVIN SCOTT:** Yeah, researcher. You know, like you were just telling me, you've got like this stack of submissions that are going into the neural --

(Crosstalk.)

**FEI-FEI LI:** -- in 24 hours.

**KEVIN SCOTT:** Yeah, and so like, thank you, by the way, for doing the podcast when --

**FEI-FEI LI:** No problem.

**KEVIN SCOTT:** -- you've got this big deadline. So, but, you know, aside from these things, like, what do you -- what do you do for fun? Like, what's --

**FEI-FEI LI:** You know what? My students asked me the same question a month ago (Laughter.) and they even laughed. They don't think I could have a good answer. And, I don't know if I could have a good answer.

So, what do you do define fun? The bliss for myself is that my work is fun to me. Hanging out with my kids is fun for me. I mean, granted, if they throw a tantrum, it's not a fun moment. But I love being with my kids. I love my students, talking to them on research ideas. Even if we come up with a bunch of stupid ideas, that process is fun.

HAI is so much fun meeting, you know, we have 200-plus faculty across the campus working on different aspects. Just talking to any one of them is fun.

So, from that point of view, I mean, I do miss some of the -- early, pre-kids, two-people world where my husband and I would go for movies or travel to a foreign country. I haven't had that for a while, that I -- I mean, I travel, but not for vacation.

But I love reading. I always read a lot of different books. I love food. Good food is always fun. (Laughter.)

Yeah, so when I was a kid, I actually do painting. One day, I will pick that up.

**KEVIN SCOTT:** Yeah. (Laughter.) Well, I think you're right, like --

**FEI-FEI LI:** What do you do for fun?

**KEVIN SCOTT:** I pray for more time because I -- (laughter)

**FEI-FEI LI:** Podcasting is fun. (Laughter.)

**KEVIN SCOTT:** This is ah… It’s really just a great thing to be able to enjoy the work that you're doing.

**FEI-FEI LI:** Yes.

**KEVIN SCOTT:** To sort of combine the things that you're most interested in.

**FEI-FEI LI:** Yeah.

**KEVIN SCOTT:** With the things that, you know, are somehow creating some sort of positive benefit for other people. And so --

**FEI-FEI LI:** Yeah, like --

(Crosstalk.)

**KEVIN SCOTT:** -- like that for me is fun --

**FEI-FEI LI:** Feeling passionate -- right, exactly, like, in one month, Stanford is going to welcome our 2019 class of AI4ALL students. I'm just so looking forward to that, right? Like, I'll be getting to know another group of 32 unbelievable high schoolers.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** That's fun. (Laughter.)

**KEVIN SCOTT:** Yeah, my fun, like, if I had to boil it down into two things, it is being able to do something that fulfills my curiosity and --

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** -- and to be able to make things. Like, those are, you know, aside from, like, my number-one fun thing is being with my kids. But, you know, like, if I'm just sort of looking selfishly at my -- at myself, it's like the -- you know, sort of the curiosity of making.

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** Yeah.

**FEI-FEI LI:** Absolutely.

**KEVIN SCOTT:** Awesome, well --

**FEI-FEI LI:** Being a professor is a lot of fun. (Laughter.)

**KEVIN SCOTT:** Yeah, which is great to hear because, like, I contemplated being a professor for a long while.

**FEI-FEI LI:** It's never too late. (Laughter.)

**KEVIN SCOTT:** No, I think it might be too late for me, Fei-Fei. (Laughter.)

**FEI-FEI LI:** Well, if you don't try, how do you know? (Laughter.)

**KEVIN SCOTT:** Well, thank you --

**FEI-FEI LI:** Okay, thank you.

**KEVIN SCOTT:** Thank you so much for being on the podcast.

**FEI-FEI LI:** Thank you, Kevin.

**KEVIN SCOTT:** And more importantly, like, thank you for, like, all of the great work that you're doing now trying to make AI more human --

**FEI-FEI LI:** And good luck to your book. I look forward to that.

**KEVIN SCOTT:** Yeah, thank you.

**FEI-FEI LI:** All right.

**KEVIN SCOTT:** Awesome.

**FEI-FEI LI:** Thank you.

**[MUSIC]**

**CHRISTINA WARREN:** We hope you enjoyed Kevin's interview with Fei-Fei Li, researcher and professor at Stanford University.

So, what I thought was really interesting about this conversation, Kevin, was she described her “convoluted,” in her words, entry point to computer science and that her passion for physics was kind of what led her into this.

But this is something that we've kind of seen with other guests on the show where people have an untraditional way of getting into these subjects.

**KEVIN SCOTT:** Yeah, and it's really amazing, like a bunch of different people find their way to AI from a bunch of different paths, although the physics one is not that uncommon.

**CHRISTINA WARREN:** I was going to say, that's the one we've kind of heard again and again is that there's something to that, I guess.

**KEVIN SCOTT:** I've been thinking for a bit about this idea that maybe artificial intelligence for human intelligence is sort of the same thing as physics is for the natural world. So, like when you think about physics, it's the way that a curious person can approach like all of these super-complicated phenomena that occur in the natural world. And so, like you can describe them and like build these models up and then understand them and be able to predict them.

And human intelligence is this super-duper complicated thing. I'm sure that one of the first thoughts that human beings had as soon as we were self-aware and had language was like, what is this thing, like, you know, why do I have the thoughts that I have, like what is -- what's the nature of my own intelligence.

And so, we've been thinking about it philosophically for thousands and thousands of years, just sort of the nature of human intelligence. And we've been thinking about it scientifically for several hundred years now with increasingly better, you know, sort of biology and neuroscience.

But still the phenomenon are so complicated that, you know, one of the things that may be very interesting about AI is that it could be a system that shines light on how human intelligence actually works by giving us a way to model it in some, you know, analytical system.

**CHRISTINA WARREN:** Yeah, no, and that's really interesting, too, I think when you look at the role that physics obviously plays with things like quantum and how that could then also go into looking at those models and furthering AI and whatnot.

And that kind of leads me to another thing that Fei-Fei was talking about and you were as well with this concept of human-centered AI and the idea that AI isn't only computer science, that it's interdisciplinary.

**KEVIN SCOTT:** Yeah. Yeah, and I think we can see that more and more all the time. I mean, it's very obvious to me at least that if you are building a technology that's going to have such a massive potential impact on the world, that you want everyone to be thinking as hard as they possibly can about how to make sure that it is sort of providing some set of human-centered benefits that are equitably distributed and sort of fair, like all of the things that we just sort of want for society itself, like we should embed into AI and guide its development accordingly.

And I think that's not just -- I mean, it's obviously not a computer science only thing, it has to be about philosophers and ethicists and economists and business folks and historians and writers and artists. And so, like we really, really have to make all of this a multidisciplinary effort if we want to get a thing that is truly a reflection of our own humanity. One of the things that I really liked about your discussion is that oftentimes you and I when we kind of have talk -- when we talk about AI, it's about like the downsides or the potential challenges and maybe the scary aspects.

In this case, really the idea of AI, you know, augmenting or enhancing and helping rather than replacing how things work in the world, how can we use this to make things better rather than how is this a threat.

**KEVIN SCOTT:** Yeah. And this is the thing that I tell people all the time, AI is just another tool that we human beings have invented to do things, and like we get to choose what we have the tool do.

And like when we make choices about say, for instance, applying AI to healthcare to make things less expensive and more accessible and higher quality for everyone, like it obviously creates this amazing positive human benefit.

And so, like I think the trick to getting, you know, the balance of AI to be beneficial and good for everyone is us choosing to do that. And so, it's like really, really amazing to have a computer scientists and one of the pioneers of the field like Fei-Fei spending so much of her energy thinking about what those beneficial applications of AI are.

**CHRISTINA WARREN:** No, I totally agree. It definitely makes me feel better, I guess, about like the future both of like humanity and, you know, the world we live in with all of this stuff.

**KEVIN SCOTT:** Yeah. I have faith in us.

**CHRISTINA WARREN:** Great. Okay, so we are out of time for now, but if you haven't listened to all of our past podcasts, you might want to spend a few minutes catching up. Now, Kevin, do you have a favorite past episode?

**KEVIN SCOTT:** Oh, I love them all equally, like my children.

**CHRISTINA WARREN:** My children are like my -- it's like my media collection. So, I know how you feel.

Oh, okay, but our listeners will have to make the decision for themselves. But you can also write to us anytime at BehindtheTech@microsoft.com, and tell us what's your favorite show and maybe what you'd like to hear more about.

**KEVIN SCOTT:** Yeah, absolutely, we'd love to hear from you. And with that, we'll see you next time!

**[MUSIC]**