

DR FEI FEI LI

LENNY'S PODCAST

BILINGUAL TRANSCRIPT

ORIGINAL BY

Lenny Rachitsky

@lennysan • x.com/lennysan

ANALYSIS BY

@Penny777 • x.com/penny777

Dr. Fei Fei Li - 双语对照

This is the complete bilingual transcript of Lenny's Podcast featuring Dr. Fei-Fei Li.

[00:00:00] Lenny Rachitsky

English:

A lot of people call you the godmother of AI. The work you did actually was the spark that brought us out of AI winter.

中文翻译:

很多人称你为“AI 教母”。你所做的工作实际上是点燃火花、带领我们走出“AI 寒冬”的关键。

[00:00:07] Dr. Fei-Fei Li

English:

In the middle of 2015, middle of 2016, some tech companies avoid using the word AI because they were not sure if AI was a dirty word. 2017-ish was the beginning of companies calling themselves AI companies.

中文翻译:

在 2015 年到 2016 年中期，一些科技公司甚至会刻意避免使用“AI”这个词，因为他们不确定 AI 是否已经变成了一个贬义词。直到 2017 年左右，公司才开始纷纷称自己为“AI 公司”。

[00:00:22] Lenny Rachitsky

English:

There's this line, I think, this was when you were presenting to Congress. There's nothing artificial about AI. It's inspired by people. It's created by people, and most importantly, it impacts people.

中文翻译:

有一句话我印象很深，我想那是你在国会作证时说的：“人工智能（AI）中没有任何东西是‘人造’的。它受人启发，由人创造，最重要的是，它影响着人。”

[00:00:30] Dr. Fei-Fei Li

English:

It's not like I think AI will have no impact on jobs or people. In fact, I believe that whatever AI does, currently or in the future, is up to us. It's up to the people. I do believe technology is a net positive for humanity, but I think every technology is a double-edged sword. If we're not doing the right thing as a society, as individuals, we can screw this up as well.

中文翻译:

我并不是认为 AI 对工作或人类没有影响。事实上，我相信无论 AI 现在或未来做什么，都取决于我们，取决于人类。我确实相信技术对人类来说是利大于弊的，但我认为每项技术都是一把双刃剑。如果我们作为一个社会、作为个人没有做正确的事，我们同样会把事情搞砸。

[00:00:56] Lenny Rachitsky

English:

You had this breakthrough insight of just, okay, we can train machines to think like humans, but it's just missing the data that humans have to learn as a child.

中文翻译:

你当时有一个突破性的洞察：我们可以训练机器像人一样思考，但它只是缺少了人类在孩童时期学习时所拥有的那种数据。

[00:01:03] Dr. Fei-Fei Li

English:

I chose to look at artificial intelligence through the lens of visual intelligence because humans are deeply visual animals. We need to train machines with as much information as possible on images of objects, but objects are very, very difficult to learn. A single object can have infinite possibilities that is shown on an image. In order to train computers with tens and thousands of object concepts, you really need to show it millions of examples.

中文翻译:

我选择通过“视觉智能”的视角来看待人工智能，因为人类是高度视觉化的动物。我们需要用尽可能多的物体图像信息来训练机器，但物体是非常、非常难以学习的。同一个物体在图像中展现出的可能性是无限的。为了让计算机训练出成千上万个物体概念，你真的需要给它展示数百万个例子。

[00:01:36] Lenny Rachitsky

English:

Today, my guest is Dr. Fei-Fei Li, who's known as the godmother of AI. Fei-Fei has been responsible for and at the center of many of the biggest breakthroughs that sparked the AI revolution that we're currently living through. She spearheaded the creation of ImageNet, which was basically her realizing that AI needed a ton of clean-labeled data to get smarter, and that data set became the breakthrough that led to the current approach to building and scaling AI models. She was chief AI scientist at Google Cloud, which is where some of the biggest early technology breakthroughs emerged from. She was director at SAIL, Stanford's Artificial Intelligence Lab, where many of the biggest AI minds came out of. She's also co-creator of Stanford's Human-Centered AI Institute, which is playing a vital role in a direction that AI is taking. She's also been on the board of Twitter. She was named one of Time's 100 Most Influential People in AI. She's also United Nations advisory board. I could go on.

(00:02:29):

In our conversation, Fei-Fei shares a brief history of how we got to today in the world of AI, including this mind-blowing reminder that 9 to 10 years ago, calling yourself an AI company was basically a death knell for your brand because no one believed that AI was actually going to work. Today, it's completely

different. Every company is an AI company. We also chat about her take on how she sees AI impacting humanity in the future, how far current technologies will take us, why she's so passionate about building a world model and what exactly world models are, and most exciting of all, the launch of the world's first large world model, Marble, which just came out as this podcast comes out. Anyone can go play with this at marble.worldlabs.ai. It's insane. Definitely check it out. Fei-Fei is incredible and way too under the radar for the impact that she's had on the world, so I am really excited to have her on and to spread her wisdom with more people.

(00:03:22):

A huge thank you to Ben Horowitz and Condoleezza Rice for suggesting topics for this conversation. If you enjoy this podcast, don't forget to subscribe and follow it in your favorite podcasting app or YouTube. With that, I bring you Dr. Fei-Fei Li after a short word from our sponsors.

中文翻译:

今天，我的嘉宾是被称为“AI 教母”的李飞飞博士 (Dr. Fei-Fei Li)。飞飞一直负责并处于许多重大突破的核心，正是这些突破引发了我们目前正在经历的 AI 革命。她主导创建了 ImageNet，这源于她意识到 AI 需要海量的清洗标注数据才能变得更聪明，而这个数据集也成为了突破口，引领了当前构建和扩展 AI 模型的方法。她曾担任 Google Cloud 的首席 AI 科学家，那里诞生了一些早期的重大技术突破。她曾是斯坦福大学人工智能实验室 (SAIL) 的主任，许多顶尖的 AI 人才都出自那里。她还是斯坦福人文 AI 研究院 (HAI) 的共同创始人，该机构在引导 AI 发展方向方面发挥着至关重要的作用。她还曾担任 Twitter 的董事，被《时代》杂志评为 AI 领域百大影响力人物之一，并且是联合国咨询委员会成员。我还可以列举更多。

(00:02:29):

在我们的对话中，飞飞分享了 AI 世界如何发展到今天的简史，包括一个令人惊讶的提醒：在 9 到 10 年前，称自己为 AI 公司基本上是品牌的“丧钟”，因为没人相信 AI 真的能成功。而今天完全不同了，每家公司都是 AI 公司。我们还聊到了她对 AI 未来如何影响人类的看法，当前技术能走多远，为什么她对构建“世界模型” (World Model) 如此充满热情，以及世界模型究竟是什么。最令人兴奋的是，世界上第一个大世界模型 Marble 的发布，它正好随本期播客同步推出。任何人都可以去 marble.worldlabs.ai 体验，非常震撼，一定要去看看。飞飞非常了不起，相对于她对世界产生的影响，她本人实在太低调了，所以我非常激动能邀请她来，并向更多人传播她的智慧。

(00:03:22):

非常感谢 Ben Horowitz 和 Condoleezza Rice 为本次对话提供的选题建议。如果你喜欢这个播客，别忘了在常用的播客应用或 YouTube 上订阅和关注。下面，在听完赞助商的简短介绍后，让我们欢迎李飞飞博士。

[00:03:37] Lenny Rachitsky (Sponsor: Figma)

English:

This episode is brought to you by Figma, makers of Figma Make. When I was a PM at Airbnb, I still remember when Figma came out and how much it improved how we operated as a team. Suddenly, I can involve my whole team in the design process, give feedback on design concepts really quickly and it just made the whole product development process so much more fun. But Figma never felt like it was for me. It was great for giving feedback and designs, but as a builder, I wanted to make stuff. That's why Figma built Figma Make. With just a few prompts, you can make any idea or design into a fully functional prototype or app that anyone can iterate on and validate with customers. Figma Make is a different kind of vibe coding tool. Because it's all in Figma, you can use your team's existing design building blocks, making it easy to create outputs that look good and feel real and are connected to how your team builds. Stop spending so much time telling people about your product vision and instead show it to them. Make code-back prototypes and apps fast with Figma Make. Check it out at figma.com/lenny.

中文翻译:

本期节目由 Figma 赞助，他们是 Figma Make 的创造者。当我在 Airbnb 担任产品经理时，我仍记得 Figma 问世时的情景，以及它如何极大地改善了我们的团队运作方式。突然间，我可以让整个团队参与到设计过程中，快速对设计概念给出反馈，这让整个产品开发过程变得有趣得多。但以前我觉得 Figma 并不完全是我准备的。它在反馈和设计方面很棒，但作为一个构建者，我想亲手做出东西。这就是为什么 Figma 开发了 Figma Make。只需几个提示词，你就可以将任何想法或设计变成功能齐全的原型或应用，任何人都可以对其进行迭代并与客户进行验证。Figma Make 是一种不同风格的“氛围编程”(vibe coding) 工具。因为它完全集成在 Figma 中，你可以使用团队现有的设计组件，轻松创建出既美观又真实、且与团队构建方式相连的产出。别再花那么多时间向别人描述你的产品愿景了，直接展示给他们看吧。使用 Figma Make 快速制作代码驱动的原型和应用。请访问 figma.com/lenny 了解详情。

[00:04:40] Lenny Rachitsky (Sponsor: Justworks)

English:

Did you know that I have a whole team that helps me with my podcast and with my newsletter? I want everyone on that team to be super happy and thrive in the roles. Justworks knows that your employees are more than just your employees; they're your people. My team is spread out across Colorado, Australia, Nepal, West Africa, and San Francisco. My life would be so incredibly complicated to hire people internationally, to pay people on time and in their local currencies, and to answer their HR questions 24/7. But with Justworks, it's super easy. Whether you're setting up your own automated payroll, offering premium benefits, or hiring internationally, Justworks offer simple software and 24/7 human support from small business experts for you and your people. They do your human resources right so that you can do right by your people. Justworks, for your people.

中文翻译:

你知道吗，我有一整个团队在帮我做播客和时事通讯。我希望团队中的每个人都能开心地工作并在岗位上发光发热。Justworks 明白，你的员工不仅仅是员工，他们是你的伙伴。我的团队分布在科罗拉多、澳大利亚、尼泊尔、西非和旧金山。如果要我亲自处理跨国招聘、按时用当地货币发放薪水、以及全天候回答人力资源问题，我的生活会变得极其复杂。但有了 Justworks，一切都变得非常简单。无论你是要设置自动化薪酬管理、提供优质福利，还是进行国际招聘，Justworks 都能为你和你的伙伴提供简单的软件支持以及来自小企业专家的 24/7 人工支持。他们帮你处理好人力资源，让你能全心对待你的伙伴。Justworks，为你的伙伴而生。

[00:05:31] Lenny Rachitsky

English:

Fei-Fei, thank you so much for being here and welcome to the podcast.

中文翻译:

飞飞，非常感谢你能来，欢迎来到本播客。

[00:05:34] Dr. Fei-Fei Li

English:

I'm excited to be here, Lenny.

中文翻译:

很高兴来到这里，Lenny。

[00:05:36] Lenny Rachitsky

English:

I'm even more excited to have you here. It is such a treat to get to chat with you. There's so much that I want to talk about. You've been at the center of this AI explosion that we're seeing right now for so long. We're going to talk about a bunch of the history that I think a lot of people don't even know about how this whole thing started, but let me first read a quote from Wired about you just so people get a sense, and in the intro I'll share all of the other epic things you've done. But I think this is a good way to just set context. "Fei-Fei is one of a tiny group of scientists, a group perhaps small enough to fit around a kitchen table, who are responsible for AI's recent remarkable advances."

(00:06:10):

A lot of people call you the godmother of AI, and unlike a lot of AI leaders, you're an AI optimist. You don't think AI is going to replace us. You don't think it's going to take all our jobs. You don't think it's going to kill us. So I thought it'd be fun to start there, just what's your perspective on how AI is going to impact humanity over time?

中文翻译:

能邀请到你我更激动。能和你聊天真是太荣幸了。我有太多想聊的话题。你长期以来一直处于我们现在看到的AI爆发的核心。我们会聊聊很多大家可能都不知道的关于这一切是如何开始的历史。但首先，我想读一段《连线》(Wired)杂志关于你的评价，让大家有个概念，在开场白中我会分享你做的其他了不起的事情。我觉得这段话很好地设定了背景：“飞飞是极少数科学家中的一员——这个群体可能小到可以围坐在一张餐桌旁——他们对AI近期取得的显著进步负有直接责任。”

(00:06:10):

很多人称你为AI教母，而且与许多AI领袖不同，你是一个AI乐观主义者。你不认为AI会取代我们，不认为它会抢走我们所有的工作，也不认为它会毁灭人类。所以我想从这里开始，你对AI随着时间的推移将如何影响人类有什么看法？

[00:06:30] Dr. Fei-Fei Li

English:

Yeah, okay, so Lenny, let me be very clear. I'm not a utopian, so it's not like I think AI will have no impact on jobs or people. In fact, I'm a humanist. I believe that whatever AI does, currently or in the future, is up to us. It's up to the people. So I do believe technology is a net positive for humanity. If you look at the long course of civilization, I think we are, and fundamentally, we're an innovative species that we... If you look at from written record thousands of years ago to now, humans just kept innovating ourselves and innovating our tools, and with that, we make lives better, we make work better, we build civilization, and I do believe AI is part of that. So that's where the optimism comes from. But I think every technology is a double-edged sword, and if we're not doing the right thing as a species, as a society, as communities, as individuals, we can screw this up as well.

中文翻译:

好的，Lenny，让我先明确一点。我不是一个乌托邦主义者，所以我并不是认为AI对工作或人类没有影响。事实上，我是一个人文主义者。我相信无论AI现在或未来做什么，都取决于我们，取决于人类。我确实相信技术对人类来说是利大于弊的。如果你回顾漫长的文明历程，我认为我们从根本上是一个创新的物种。从几千年前的有文字记录到现在，人类一直在不断地自我创新和工具创新，通过这些，我们让生活变得更好，让工作变得

更好，建立了文明，而我相信 AI 也是其中的一部分。这就是乐观主义的来源。但我认为每项技术都是一把双刃剑，如果我们作为一个物种、一个社会、一个社区或个人没有做正确的事，我们同样会把事情搞砸。

[00:07:47] Lenny Rachitsky

English:

There's this line, I think, this was when you were presenting to Congress, "There's nothing artificial about AI. It's inspired by people. It's created by people, and most importantly, it impacts people." I don't have a question there, but what a great line.

中文翻译:

我想起那句话，应该是你在国会作证时说的：“人工智能中没有任何东西是‘人造’的。它受人启发，由人创造，最重要的是，它影响着人。”我这里没有具体问题，只是觉得这句话说得太棒了。

[00:07:59] Dr. Fei-Fei Li

English:

Yeah, I feel pretty deeply. I started working AI two and a half decades ago, and I've been having students for the past two decades and almost every student who graduates, I remind them when they graduate from my lab that your field is called artificial intelligence, but there's nothing artificial about it.

中文翻译:

是的，我感触很深。我从 25 年前就开始从事 AI 工作，过去 20 年里我带过很多学生，几乎每个从我实验室毕业的学生，我都会提醒他们：你们的领域虽然叫“人工智能”，但它没有任何东西是“人造”的（强调其背后的人文属性）。

[00:08:23] Lenny Rachitsky

English:

Coming back to the point you just made about how it's kind of up to us about where this all goes, what is it you think we need to get right? How do we set things on a path? I know this is a very difficult question to answer, but just what's your advice? What do you think we should be keeping in mind?

中文翻译:

回到你刚才说的，这一切的发展方向取决于我们。你认为我们需要做对哪些事情？我们该如何设定正确的发展路径？我知道这是一个很难回答的问题，但你的建议是什么？你认为我们应该记住什么？

[00:08:36] Dr. Fei-Fei Li

English:

Yeah, how many hours do we have?

中文翻译:

额，我们有几个小时的时间来聊这个？

[00:08:39] Lenny Rachitsky

English:

How do we align AI? There we go. Let's solve it.

中文翻译:

我们如何实现“AI 对齐”（Align AI）？来吧，让我们解决它。

[00:08:41] Dr. Fei-Fei Li

English:

So I think people should be responsible individuals no matter what we do. This is what we teach our children, and this is what we need to do as grownups as well. No matter which part of the AI development or AI deployment or AI application you are participating in, and most likely many of us, especially as technologists, we're in multiple points. We should act like responsible individuals and care about this. Actually, care a lot about this. I think everybody today should care about AI because it is going to impact your individual life. It is going to impact your community, it's going to impact the society and the future generation. And caring about it as a responsible person is the first, but also the most important step.

中文翻译:

我认为无论我们做什么，都应该做一个负责任的个体。这是我们教给孩子的东西，也是我们作为成年人需要做的。无论你参与的是 AI 开发、部署还是应用的哪个环节——而且很可能我们中的许多人，尤其是技术人员，身兼数职——我们都应该表现得像个负责任的人，并去关心这件事。实际上，要非常关心。我认为当今每个人都应该关心 AI，因为它将影响你的个人生活、你的社区、整个社会以及下一代。作为一个负责任的人去关心它，是第一步，也是最重要的一步。

[00:09:37] Lenny Rachitsky

English:

Okay, so let me actually take a step back and kind of go to the beginning of AI. Most people started hearing and caring about AI, as what it's called today, just like, I don't know, a few years ago when ChatGPT came out. Maybe it was like three years ago.

中文翻译:

好，让我退后一步，回到 AI 的起点。大多数人开始听说并关心现在所谓的“AI”，大概也就是几年前 ChatGPT 问世的时候。也许是三年前？

[00:09:51] Dr. Fei-Fei Li

English:

Three years ago, almost one more month, three years ago.

中文翻译:

三年前，再过一个月就满三年了。

[00:09:55] Lenny Rachitsky

English:

Wow, okay. And that was ChatGPT coming out. Is that the milestone you have in mind?

中文翻译:

哇，好的。那就是 ChatGPT 的发布。那是你心目中的里程碑吗？

[00:09:56] Dr. Fei-Fei Li

English:

Yes.

中文翻译:

是的。

[00:09:57] Lenny Rachitsky

English:

Okay, cool. That's exactly how I saw it. But very few people know there was a long, long history of people working on, it was called machine learning back then and there's other terms, and now it's just everything's AI and there was kind of a long period of just a lot of people working on it. And then there's this what people refer to as the AI winter where people just gave up almost, most people did, and just, okay, this idea isn't going anywhere. And then the work you did actually was essentially the spark that brought us out of AI winter and is directly responsible for the world where now of just AI is all we talk about. As you just said, it's going to impact everything we do. So I thought it'd be really interesting to hear from you just the brief history of what the world was like before ImageNet and just the work you did to create ImageNet, why that was so important, and then just what happened after.

中文翻译:

好的，太酷了。我也是这么看的。但很少有人知道，在此之前有一段非常漫长的历史，人们当时称之为“机器学习”（Machine Learning）或其他术语，而现在一切都叫 AI。曾有一段很长的时间，很多人在研究它，然后出现了人们所说的“AI 寒冬”，那时人们几乎放弃了，大多数人觉得这个想法行不通。而你所做的工作实际上是带领我们走出 AI 寒冬的火花，并直接造就了我们现在这个开口闭口都是 AI 的世界。正如你刚才所说，它将影响我们所做的一切。所以我觉得听你讲讲 ImageNet 之前世界是什么样子的，你为创建 ImageNet 做了哪些工作，为什么它如此重要，以及之后发生了什么，会非常有趣。

[00:10:44] Dr. Fei-Fei Li

English:

It is, for me, hard to keep in mind that AI is so new for everybody when I lived my entire professional life in AI. There's a part of me that is just, it's so satisfying to see a personal curiosity that I started barely out of teenagehood and now has become a transformative force of our civilization. It generally is a civilizational level technology. So that journey is about 30 years or 20 something, 20 plus years, and it's just very satisfying. So where did it all start? Well, I'm not even the first generation AI researcher. The first generation really date back to the '50s and '60s, and Alan Turing was ahead of his time in the '40s by asking, daring humanity with the question, "Is there thinking machines?" And of course he has a specific way of testing this concept of thinking machine, which is a conversational chatbot, which to his standard we now have a thinking machine.

(00:12:02):

But that was just a more anecdotal inspiration. The field really began in the '50s when computer scientists came together and look at how we can use computer programs and algorithms to build these programs that can do things that have been only capable by human cognition. And that was the beginning. And the founding fathers the Dartmouth workshop in the 1956, we have Professor John McCarthy who later came to Stanford who coined the term artificial intelligence. And between the '50s, '60s, '70s, and '80s, it was the early days of AI exploration and we had logic systems, we had expert systems, we also had early exploration of neural network. And then it came to around the late '80s, the '90s, and the very beginning of the 21st century. That stretch about 20 years is actually the beginning of machine learning, is the marriage between computer programming and statistical learning.

(00:13:23):

And that marriage brought a very, very critical concept into AI, which is that purely rule-based program is not going to account for the vast amount of cognitive capabilities that we imagine computers can do. So we have to use machines to learn the patterns. Once the machines can learn the patterns, it has a hope to do more things. For example, if you give it three cats, the hope is not just for the machines to recognize these three cats. The hope is the machines can recognize the fourth cat, the fifth cat, the sixth cat, and all the other cats. And that's a learning ability that is fundamental to humans and remaining animals. And we, as a field, realized, "We need machine learning." So that was up till the beginning of the 21st century. I entered the field of AI literally in the year of 2000. That's when my PhD began at Caltech.

(00:14:33):

And so I was one of the first generation machine learning researchers and we were already studying this concept of machine learning, especially neural network. I remember that was one of my first courses at Caltech is called neural network, but it was very painful. It was still smack in the middle of the so-called AI winter, meaning the public didn't look at this too much. There wasn't that much funding, but there was also a lot of ideas flowing around. And I think two things happened to myself that brought my own career so close to the birth of modern AI is that I chose to look at artificial intelligence through the lens of visual intelligence because humans are deeply visual animals. We can talk a little more later, but so much of our intelligence is built upon visual, perceptual, spatial understanding, not just language per se. I think they're complementary.

(00:15:37):

So I choose to look at visual intelligence and my PhD and my early professor years, my students and I are very committed to a north star problem, which is solving the problem of object recognition because it's a building block for the perceptual world, right? We go around the world interpreting reasoning and interacting with it more or less at the object level. We don't interact with the world at the molecular level. We don't interact with the world as... We sometimes do, but we rarely, for example, if you want to lift a teapot, you don't say, "Okay, the teapot is made of a hundred pieces of porcelain and let me work on this a hundred pieces." You look at this as one object and interact with it. So object is really important. So I was among the first researchers to identify this as a north star problem, but I think what happened is that as a student of AI and a researcher of AI, I was working on all kinds of mathematical models including neural network, including Bayesian network, including many, many models.

(00:16:53):

And there was one singular pain point is that these models don't have data to be trained on. And as a field, we were so focusing on these models, but it dawned on me that human learning as well as evolution is actually a big data learning process. Humans learn with so much experience constantly. In the evolution, if you look at time, animals evolve with just experiencing the world. So I think my students

and I conjectured that a very critically-overlooked ingredient of bringing AI to life is big data. And then we began this ImageNet project in 2006, 2007. We were very ambitious. We want to get the entire internet's image data on objects. Now granted internet was a lot smaller than today, so I felt like that ambition was at least not too crazy. Now, it's totally delusional to think a couple of graduate student and a professor can do this.

(00:18:05):

And that's what we did. We curated very carefully, 15 million images on the internet, created a taxonomy of 22,000 concepts, borrowing other researchers' work like linguists work on WordNet, and it's a particular way of dictionarying words. And we combine that into ImageNet and we open-sourced that to the research community. We held an annual ImageNet challenge to encourage everybody to participate in this. We continue to do our own research, but 2012 was the moment that many people think was the beginning of the deep learning or birth of modern AI because a group of Toronto researchers led by Professor Geoff Hinton, participated in ImageNet Challenge, used ImageNet big data and two GPUs from NVIDIA and created successfully the first neural network algorithm that can...

(00:19:12):

It didn't totally solve, but made a huge progress towards solving the problem of object recognition. And that combination of the trio technology, big data, neural network, and GPU was kind of the golden recipe for modern AI. And then fast-forward, the public moment of AI, which is the ChatGPT moment, if you look at the ingredients of what brought ChatGPT to the world technically still use these three ingredients. Now, it's internet-scale data mostly texts is a much more complex neural network architecture than 2012, but it's still neural network and a lot more GPUs, but it's still GPUs. So these three ingredients are still at the core of modern AI.

中文翻译:

对我来说，很难时刻记住 AI 对大家来说是如此新鲜的事物，因为我的整个职业生涯都沉浸在 AI 中。看到一个我从青少年时期就开始的个人好奇心，现在变成了改变文明的力量，我感到非常欣慰。它确实是一项文明级别的技术。这段旅程大约持续了 30 年，或者说 20 多年，这非常令人满足。那么一切是从哪里开始的呢？其实，我甚至不算第一代 AI 研究者。第一代真的要追溯到 50 和 60 年代。艾伦·图灵 (Alan Turing) 在 40 年代就超前于时代，向人类提出了一个大胆的问题：“会有会思考的机器吗？”当然，他有一种特定的测试这种“思考机器”概念的方法，那就是对话式聊天机器人——按照他的标准，我们现在已经拥有了会思考的机器。

(00:12:02):

但那更多只是轶事般的启发。这个领域真正开始于 50 年代，当时计算机科学家聚集在一起，研究如何利用计算机程序和算法来构建能够完成只有人类认知才能完成的任务的程序。那是开端。1956 年达特茅斯会议的创始人之一，后来来到斯坦福的约翰·麦卡锡 (John McCarthy) 教授，创造了“人工智能”这个词。在 50、60、70 和 80 年代之间，是 AI 探索的早期阶段，我们有了逻辑系统、专家系统，也有了对神经网络的早期探索。然后到了 80 年代后期、90 年代以及 21 世纪初。那段大约 20 年的时间实际上是机器学习的开端，是计算机编程与统计学习的结合。

(00:13:23):

这种结合为 AI 引入了一个非常关键的概念，即纯粹基于规则的程序无法涵盖我们设想中计算机能做的海量认知功能。因此，我们必须让机器去学习模式。一旦机器能学习模式，它就有希望做更多事情。例如，如果你给它看三只猫，目标不仅仅是让机器识别这三只猫，而是希望它能识别第四只、第五只、第六只以及所有其他的猫。这种学习能力是人类和其他动物的基础。我们作为一个领域意识到：“我们需要机器学习。”这种情况一直持续到 21 世纪初。我是在 2000 年正式进入 AI 领域的，那是我在加州理工学院 (Caltech) 开始读博的时候。

(00:14:33):

所以我是第一代机器学习研究者之一，我们当时已经在研究机器学习的概念，尤其是神经网络。我记得我在加州理工的第一门课就叫“神经网络”，但当时非常痛苦。那时仍处于所谓的“AI 寒冬”正中心，意味着公众不太关注，资金也不多，但也有很多想法在涌动。我认为有两件事发生在我身上，让我的职业生涯如此接近现代 AI 的诞生：第一，我选择通过视觉智能的视角来看待人工智能，因为人类是深度视觉化的动物。我们稍后可以多聊聊，但我们的智能很大程度上建立在视觉、感知和空间理解之上，而不仅仅是语言本身。我认为它们是互补的。

(00:15:37):

所以我选择研究视觉智能。在我的博士阶段和早期的教授生涯中，我和我的学生们致力于一个“北极星问题”，即解决物体识别问题，因为它是感知世界的基石。我们在世界上走动，解释、推理并与之互动，基本上都是在物体层面进行的。我们不会在分子层面与世界互动。例如，如果你想拿起一个茶壶，你不会说：“好的，这个茶壶由一百块瓷片组成，让我处理这一百块瓷片。”你会把它看作一个整体并与之互动。所以物体非常重要。我是最早将此确定为核心问题的研究者之一。但后来发生的是，作为 AI 的学生和研究者，我当时在研究各种数学模型，包括神经网络、贝叶斯网络等许多模型。

(00:16:53):

当时有一个唯一的痛点：这些模型没有数据可以训练。作为一个领域，我们当时太专注于模型了，但我突然意识到，人类的学习以及进化实际上是一个大数据学习过程。人类不断地通过海量经验进行学习。在进化过程中，动物通过体验世界而进化。所以我和我的学生推测，赋予 AI 生命的一个被严重忽视的关键要素是“大数据”。于是我们在 2006、2007 年开始了 ImageNet 项目。我们非常有野心，想要获取整个互联网上关于物体的图像数据。当然，当时的互联网比现在小得多，所以我觉得这个野心至少不算太疯狂。虽然现在回想起来，认为几个研究生和一个教授就能完成这件事简直是痴人说梦。

(00:18:05):

但我们确实做到了。我们非常仔细地整理了互联网上的 1500 万张图像，创建了一个包含 22,000 个概念的分类体系，借鉴了语言学家在 WordNet 上的工作（这是一种特定的单词字典编撰方式）。我们将这些结合到 ImageNet 中，并向研究界开源。我们举办了一年一度的 ImageNet 挑战赛，鼓励大家参与。我们继续自己的研究，但 2012 年是许多人认为的深度学习开端或现代 AI 诞生的时刻，因为由 Geoff Hinton 教授领导的多伦多研究小组参加了 ImageNet 挑战赛，使用了 ImageNet 大数据和两块来自 NVIDIA 的 GPU，成功创建了第一个神经网络算法，它……

(00:19:12):

它虽然没有完全解决，但在解决物体识别问题上取得了巨大进步。大数据、神经网络和 GPU 这三位一体的技术组合，成了现代 AI 的“黄金配方”。快进到 AI 的公众时刻，也就是 ChatGPT 时刻，如果你看促成 ChatGPT 问世的技术成分，依然是这三样。现在是互联网规模的数据（主要是文本），神经网络架构比 2012 年复杂得多，但它依然是神经网络，GPU 更多了，但依然是 GPU。所以这三个要素依然是现代 AI 的核心。

[00:20:16] Lenny Rachitsky

English:

Incredible. I have never heard that full story before. I love that it was two GPUs was the first. I love that. And now it's, I don't know, hundreds of thousands, right, that are orders of magnitude more powerful.

中文翻译:

太不可思议了。我以前从未听过完整的故事。我特别喜欢“最初只有两块 GPU”这个细节。而现在，天知道，是成千上万块 GPU，而且性能强了几个数量级。

[00:20:30] Dr. Fei-Fei Li

English:

Yep.

中文翻译:

没错。

[00:20:31] Lenny Rachitsky

English:

And those two GPUs where they just bought, they were like gaming GPUs, they just went to the-

中文翻译:

那两块 GPU 是买来的吗？它们就像是游戏显卡，他们直接去了……

[00:20:34] Dr. Fei-Fei Li

English:

Yes.

中文翻译:

是的。

[00:20:35] Lenny Rachitsky

English:

... GameStar that people use for playing games. As you said, this continues to be in a large way, the way models get smarter. Some of the fastest growing companies in the world right now, I've had them all mostly on the podcast, Mercor and Surge and Scale. They continue to do this for labs, just give them more and more label data of the things they're most excited and interested in.

中文翻译:

……人们用来玩游戏的显卡店。正如你所说，这在很大程度上仍然是模型变聪明的方式。目前世界上一些增长最快的公司——我几乎都请他们上过播客，比如 Mercor、Surge 和 Scale——他们继续为实验室做这件事，只是为他们提供越来越多关于他们最感兴趣领域的标注数据。

[00:20:53] Dr. Fei-Fei Li

English:

Yeah, I remember Alex Wang from Scale very early days. I probably still has his emails when he was starting Scale. He was very kind. He keeps sending me emails about how ImageNet inspired Scale. I was very pleased to see that.

中文翻译:

是的，我记得 Scale 的 Alex Wang 早期的时候。我可能还留着他刚创办 Scale 时的邮件。他非常客气，一直给我发邮件说 ImageNet 如何启发了 Scale。看到这些我非常高兴。

[00:21:08] Lenny Rachitsky

English:

One of my other favorite takeaways from what you just shared is just such an example of high agency and just doing things that's kind of a meme on Twitter. Just you can just do things. You're just like, okay, this is probably necessary to move AI. And it's called machine learning back then, right? Was that the term most people used?

中文翻译:

从你刚才的分享中，我最喜欢的另一个收获是，这展现了极强的行动力（high agency），就像Twitter上的一个梗说的：“你尽管去做”。你当时就像是觉得：好吧，为了推动AI发展，这可能是必须做的。当时它叫“机器学习”，对吧？那是大多数人使用的术语吗？

[00:21:25] Dr. Fei-Fei Li

English:

I think it was interchangeably. It's true. I do remember the companies, the tech companies, I am not going to name names, but I was in a conversation in one of the early days, I think is in the middle of 2015, middle of 2016, some tech companies avoid using the word AI because they were not sure if AI was a dirty word. And I remember I was actually encouraging everybody to use the word AI because to me that is one of the most audacious question humanity has ever asked in our quest for science and technology, and I feel very proud of this term. But yes, at the beginning some people were not sure.

中文翻译:

我想这两个词当时是混用的。确实如此。我记得那些科技公司——我就不点名了——但在早期的一次谈话中，大概是2015年中期到2016年中期，一些科技公司刻意避免使用“AI”这个词，因为他们不确定AI是否是个贬义词。我记得我当时其实是在鼓励大家使用“AI”这个词，因为对我来说，这是人类在追求科学技术过程中提出的最大胆的问题之一，我对这个词感到非常自豪。但没错，刚开始有些人确实不确定。

[00:22:12] Lenny Rachitsky

English:

What year was that roughly when AI was a dirty word?

中文翻译:

AI被视为贬义词大概是在哪一年？

[00:22:14] Dr. Fei-Fei Li

English:

2016, I think because that was-

中文翻译:

我想是2016年，因为那是……

[00:22:15] Lenny Rachitsky

English:

2016, less than 10 years ago.

中文翻译:

2016年，不到10年前。

[00:22:18] Dr. Fei-Fei Li

English:

That was the changing. Some people start calling it AI, but I think if you look at the Silicon Valley tech companies, if you trace their marketing term, I think 2017-ish was the beginning of companies calling themselves AI companies.

中文翻译:

那是转折点。有些人开始称之为AI，但如果你观察硅谷科技公司，追踪他们的营销术语，我认为2017年左右是公司开始自称“AI公司”的开端。

[00:22:40] Lenny Rachitsky

English:

That's incredible. Just how the world has changed.

中文翻译:

太不可思议了。世界变化真快。

[00:22:43] Lenny Rachitsky

English:

Now, you can't not call yourself an AI company.

中文翻译:

现在，你不能不称自己为AI公司。

[00:22:46] Dr. Fei-Fei Li

English:

I know.

中文翻译:

我知道。

[00:22:46] Lenny Rachitsky

English:

Just nine-ish years later.

中文翻译:

仅仅过了九年左右。

[00:22:48] Dr. Fei-Fei Li

English:

Yeah.

中文翻译:

是的。

[00:22:49] Lenny Rachitsky

English:

Oh, man. Okay. Is there anything else around the history, that early history that you think people don't know that you think is important before we chat about where you think things are going and the work that you're doing?

中文翻译:

天呐。好的。在聊聊你对未来的看法以及你正在做的工作之前，关于那段早期历史，还有什么你觉得大家不知道但很重要的事吗？

[00:23:01] Dr. Fei-Fei Li

English:

I think as all histories, I'm keenly aware that I am recognized for being part of the history, but there are so many heroes and so many researchers. We're talking about generations of researchers. In my own world, there are so many people who have inspired me, which I talked about in my book, but I do feel our culture, especially Silicon Valley, tends to assign achievements to a single person. While I think it has value, but it's just to be remembered. AI is a field of, at this point, 70 years old and we have gone through many generations. Nobody, no one could have gotten here by themselves.

中文翻译:

我认为就像所有历史一样，我敏锐地意识到我因为成为历史的一部分而获得认可，但其实有太多的英雄和研究者。我们谈论的是几代研究者。在我的世界里，有很多人启发了我，我在书中也提到过。但我确实觉得我们的文化，尤其是硅谷，倾向于将成就归功于某一个人。虽然我认为这有其价值，但必须记住：AI 是一个已经有 70 年历史的领域，我们经历了许多代人。没有人，绝对没有任何一个人能凭一己之力走到今天。

[00:23:54] Lenny Rachitsky

English:

Okay, so let me ask you this question. It feels like we're always on this precipice of AGI, this kind of vague term people throw around, AGI is coming, it's going to take over everything. What's your take on how far you think we might be from AGI? Do you think we're going to get there on the current trajectory we're on? Do you think we need more breakthroughs? Do you think the current approach will get us there?

中文翻译:

好，那我想问你这个问题。感觉我们总是处于 AGI（通用人工智能）的边缘，人们经常抛出这个模糊的术语，说 AGI 要来了，它将接管一切。你认为我们离 AGI 还有多远？你认为按照目前的轨迹我们能达到那里吗？你认为我们需要更多的突破吗？还是说目前的方法就能带我们到达终点？

[00:24:13] Dr. Fei-Fei Li

English:

Yeah, this is a very interesting term, Lenny. I don't know if anyone has ever defined AGI. There are many different definitions, including some kind of superpower for machines all the way to machines can become economically viable agent in the society. In other words, making salaries to live. Is that the definition of AGI? As a scientist, I take science very seriously and I enter the field because I was inspired by this audacious question of, can machines think and do things in the way that humans can do? For me, that's always the north star of AI. And from that point of view, I don't know what's the difference between AI and AGI.

(00:25:10):

I think we've done very well in achieving parts of the goal, including conversational AI, but I don't think we have completely conquered all the goals of AI. And I think our founding fathers, Alan Turing, I wonder if Alan Turing is around today and you ask him to contrast AI versus AGI, he might just shrugged and said, "Well, I asked the same question back in 1940s," so I don't want to get onto a rabbit hole of defining AI versus AGI. I feel AGI is more a marketing term than a scientific term as a scientist than technologist. AI is my north star, is my field's north star, and I'm happy people call it whatever name they want to call it.

中文翻译:

这是一个非常有趣的术语，Lenny。我不知道是否有人真正定义过 AGI。有很多不同的定义，从机器拥有某种超能力，到机器能成为社会中经济上可行的主体（换句话说，能领工资生活）。那是 AGI 的定义吗？作为一个科学家，我非常严肃地对待科学，我进入这个领域是因为我被那个大胆的问题所启发：机器能像人类一样思考和做事吗？对我来说，这永远是 AI 的北极星。从这个角度来看，我不知道 AI 和 AGI 有什么区别。

(00:25:10):

我认为我们在实现部分目标方面做得非常好，包括对话式 AI，但我认为我们还没有完全攻克 AI 的所有目标。我想，如果我们的创始人艾伦·图灵今天还在世，你让他对比 AI 和 AGI，他可能只是耸耸肩说：“好吧，我在 40 年代就问过同样的问题。”所以我不想陷入定义 AI 与 AGI 的“兔子洞”。作为一名科学家和技术专家，我觉得 AGI 更多是一个营销术语而非科学术语。AI 就是我的北极星，是我们领域的北极星，人们想怎么称呼它我都乐意。

[00:26:44] Lenny Rachitsky

English:

So let me ask you maybe this way, like you described, there's kind of these components that from ImageNet and AlexNet took us to where we're today, GPUs essentially, data, label data, just like the algorithm of the model. There's also just the transformer feels like an important step in that trajectory. Do you feel like those are the same components that'll get us to, I don't know, 10 times smarter model, something that's like life-changing for the entire world? Or do you think we need more breakthroughs? I know we're going to talk about world models, which I think is a component of this, but is there anything

else that you think is like, oh, this will plateau, or okay, this will take us just need more data, more compute, more GPUs?

中文翻译:

那也许我换个方式问。正如你所描述的，从 ImageNet 和 AlexNet 到今天，有一些核心组件：GPU、数据、标注数据，以及模型算法。Transformer 架构似乎也是这一轨迹中的重要一步。你觉得这些相同的组件能带我们走向比现在聪明 10 倍的模型，那种能改变全世界生活的模型吗？还是说我们需要更多的突破？我知道我们要聊“世界模型”，我认为那是其中的一部分，但还有什么让你觉得“哦，这会遇到瓶颈”，或者“这只需要更多数据、更多算力和更多 GPU 就能解决”？

[00:26:44] Dr. Fei-Fei Li

English:

Oh no, I definitely think we need more innovations. I think scaling loss of more data, more GPUs, and bigger current model architecture is there's still a lot to be done there, but I absolutely think we need to innovate more. There's not a single deeply scientific discipline in human history that has arrived at a place that says we're done, we're done innovating and AI is one of the, if not the youngest discipline in human civilization in terms of science and technology, we're still scratching the surface. For example, like I said, we're going to segue into world models. Today, you take a model and run it through a video of a couple of office rooms and ask the model to count the number of chairs. And this is something a toddler could do or maybe an elementary school kid could do, and AI could not do that, right?

(00:27:50):

So there's just so much AI today could not do, then let alone thinking about how did someone like Isaac Newton look at the movements of the celestial bodies and derive an equation or a set of equations that governs the movement of all bodies, that level of creativity, extrapolation, abstraction. We have no way of enabling AI to do that today. And then let's look at emotional intelligence. If you look at a student coming to a teacher's office and have a conversation about motivation, passion, what to learn, what's the problem that's really bothering you. That conversation, as powerful as today's conversational bots are, you don't get that level of emotional cognitive intelligence from today's AI. So there's a lot we can do better, and I do not believe we're done innovating.

中文翻译:

噢不，我绝对认为我们需要更多创新。虽然在更多数据、更多 GPU 和更大的现有模型架构上，缩放法则 (scaling laws) 仍有很大发挥空间，但我坚信我们需要更多创新。人类历史上没有任何一个深刻的科学学科曾到达过一个点说“我们完成了，创新结束了”。AI 即使不是人类文明中最年轻的科技学科，也是其中之一，我们还只是触及皮毛。例如，正如我所说，我们要转入世界模型的话题。今天，你拿一个模型去跑一段几个办公室的视频，让模型数一下有多少把椅子。这是一个蹒跚学步的孩子或小学生就能做到的事，但 AI 做不到，对吧？

(00:27:50):

所以今天 AI 还有太多做不到的事。更不用说像艾萨克·牛顿那样，观察天体运动并推导出统治所有物体运动的方程或方程组，那种水平的创造力、外推能力和抽象能力，我们今天根本无法让 AI 做到。再看看情感智能。如果你看到一个学生走进老师办公室，谈论动力、热情、学什么，以及真正困扰他的问题。尽管今天的对话机器人很强大，但你无法从今天的 AI 那里获得那种水平的情感能力。所以我们还有很多可以做得更好的地方，我不相信创新已经结束。

[00:29:51] Lenny Rachitsky

English:

Demis had this really interesting interview recently from DeepMind slash Google where someone asked him just like, "What do you think, how far are we from AGI? What does it look like going through there?" He had a really interesting way of approaching it is if we were to give the most cutting-edge model all the information until the end of the 20th century, see if it could come up with all the breakthroughs Einstein had and so far we're nowhere near that, but they could just-

中文翻译:

Demis (DeepMind/Google 负责人) 最近有一个非常有趣的采访，有人问他：“你认为我们离 AGI 还有多远？通往那里的过程是什么样的？”他有一个非常有趣的切入点：如果我们给最尖端的模型提供 20 世纪末之前的所有信息，看它是否能想出爱因斯坦取得的所有突破。到目前为止，我们还差得很远，但他们可以……

[00:29:22] Dr. Fei-Fei Li

English:

No, we're not. In fact, it's even worse. Let's give AI all the data including modern instruments data of celestial bodies, which Newton did not have, and give it to that and just ask AI to create the 17th century set of equations on the laws of bodily movements. Today's AI cannot do that.

中文翻译:

不，我们做不到。事实上，情况更糟。即使我们给 AI 所有数据，包括牛顿当时没有的现代仪器观测到的天体数据，然后让 AI 去创造 17 世纪那套关于物体运动规律的方程组，今天的 AI 也做不到。

[00:29:49] Lenny Rachitsky

English:

All right. We're ways away is what I'm hearing.

中文翻译:

好吧，我听明白了，我们还有很长的路要走。

[00:29:50] Dr. Fei-Fei Li

English:

Yeah.

中文翻译:

是的。

[00:29:51] Lenny Rachitsky

English:

Okay, so let's talk about world models. To me, this is just another really amazing example of you being ahead of where people end up. So you were way ahead on, okay, we just need a lot of clean data for AI and neural networks to learn. You've been talking about this idea of world models for a long time. You started a company to build, essentially there's language models. This is a different thing. This is a world

model. We'll talk about what that is. And now, as I was preparing for this Elon's talking about world models, Jensen's talking about world models, I know Google's working on this stuff. You've been at this for a long time and you actually just launched something that's going, we're going to talk about right before this podcast airs. Talk about what is a world model? Why is it so important?

中文翻译:

好，那我们聊聊“世界模型”。对我来说，这是你领先于大众认知的又一个绝佳例子。在“AI 和神经网络学习需要大量清洗数据”这一点上，你曾遥遥领先。而关于“世界模型”的想法，你也谈论很久了。你创办了一家公司来构建它。基本上，现在有语言模型，但这是不同的东西，这是世界模型。我们会聊聊它是什么。现在，当我准备这期节目时，埃隆（马斯克）在谈论世界模型，黄仁勋也在谈论世界模型，我知道 Google 也在研究。你已经研究很久了，而且就在这期播客播出前，你刚刚发布了一些东西。聊聊什么是世界模型？为什么它如此重要？

[00:30:33] Dr. Fei-Fei Li

English:

I'm very excited to see that more and more people are talking about world models like Elon, like Jensen. I have been thinking about really how to push AI forward all my life and the large language models that came out of the research world and then OpenAI and all this, for the past few years, were extremely inspiring even for a researcher like me. I remembered when GPT2 came out, and that was in, I think, late 2020. I was co-director, I still am, but I was at that time full-time co-director of Stanford's Human-Centered AI institute, and I remember it was... The public was not aware of the power of the large language model yet, but as researchers, we were seeing it, we're seeing the future, and I had pretty long conversations with my natural language processing colleagues like Percy Liang and Chris Manning. We were talking about how critical this technology is going to be and the Stanford AI Institute, Human-Centered AI Institute, HAI, was the first one to establish a full research center foundation model.

(00:31:59):

We were, Percy Liang, and many researchers led the first academic paper foundation model. So it was just very inspiring for me. Of course, I come from the world of visual intelligence and I was just thinking there's so much we can push forward beyond language because humans, humans use our sense of spatial intelligence, a world understanding to do so many things and they are beyond language. Think about a very chaotic first responder scene, whether it's fire or some traffic accident or some natural disaster. And if you immerse yourself in those scene and think about how people organize themselves to rescue people, to stop further disasters, to put down fires, a lot of that is spontaneous understanding of objects, worlds, human situational awareness. Language is part of that, but a lot of those situations, language cannot get you to put down the fire.

(00:33:21):

So that is, what is that? I was thinking a lot. And in the meantime, I was doing a lot of robotics research and it dawned on me that the linchpin of connecting the additional intelligence, in addition to language embodied AI, which are robotics, connecting visual intelligence, is the sense of spatial intelligence about understanding the world. And that's when I think it was 2024, I gave a TED talk about spatial intelligence at world models. And I start formulating this idea back in 2022 based on my robotics and computer vision research. And then one thing that was really clear to me is that I really want to work with the brightest technologists and move as fast as possible to bring this technology to life. And that's when we founded this company called World Labs. And you can see the word world is in the title of our company because we believe so much in world modeling and spatial intelligence.

中文翻译:

看到越来越多的人像埃隆、仁勋那样谈论世界模型，我非常兴奋。我一生都在思考如何推动 AI 向前发展。过去几年，从研究界、OpenAI 等诞生的语言大模型，即使对我这样的研究者来说也极具启发性。我记得 GPT-2 问世时，大概是 2020 年底。当时我是斯坦福人文 AI 研究院 (HAI) 的全职共同院长（现在依然是），我记得当时公众还没意识到语言大模型的威力，但作为研究者，我们已经看到了未来。我和我的自然语言处理同事，比如 Percy Liang 和 Chris Manning，进行了很长时间的交流。我们讨论了这项技术将变得多么关键。斯坦福 HAI 是第一个建立完整基础模型研究中心的机构。

(00:31:59):

Percy Liang 和许多研究者领导撰写了第一篇关于基础模型 (Foundation Model) 的学术论文。这对我非常有启发。当然，我来自视觉智能领域，我一直在想，除了语言之外，我们还有很多可以推进的地方。因为人类利用空间智能和对世界的理解来做很多事情，而这些是超越语言的。想象一个非常混乱的急救现场，无论是火灾、交通事故还是自然灾害。如果你置身其中，思考人们如何组织起来营救他人、阻止灾难扩大、灭火，这其中很大一部分是动作，是对物体、世界和人类情境感知的自发理解。语言是其中的一部分，但在很多情况下，语言无法帮你灭火。

(00:33:21):

那么，那是什么呢？我思考了很多。与此同时，我做了很多机器人研究，我突然意识到，将语言之外的其他智能——即具身智能 (Embodied AI，如机器人) ——与视觉智能连接起来的关键纽带，就是关于理解世界的“空间智能”。我想是在 2024 年，我做了一个关于空间智能和世界模型的 TED 演讲。基于我的机器人和计算机视觉研究，我从 2022 年就开始构思这个想法。当时我非常清楚的一点是，我真的很想与最聪明的技术专家合作，并尽可能快地将这项技术变为现实。这就是我们成立 World Labs 这家公司的初衷。你可以看到我们公司的名字里就有“World”（世界）这个词，因为我们深信世界建模和空间智能的力量。

[00:34:41] Lenny Rachitsky

English:

People are so used to just chatbots and that's a large language model. A simple way to understand a world model is you basically describe a scene and it generates an infinitely explorable world. We'll link to the thing you launched, which we'll talk about, but just is that a simple way to understand it?

中文翻译:

人们已经习惯了聊天机器人，那是语言大模型。理解世界模型的一个简单方法是：你描述一个场景，它就会生成一个可以无限探索的世界。我们会附上你发布的产品链接，稍后会细聊，但这是不是一种简单的理解方式？

[00:34:56] Dr. Fei-Fei Li

English:

That's part of it, Lenny. I think a simple way to understand a world model is that this model can allow anyone to create any worlds in their mind's eye by prompting whether it's an image or a sentence. And also be able to interact in this world whether you are browsing and walking or picking objects up or changing things as well as to reason within this world, for example, if the person consuming, if the agent consuming this output of the world model is a robot, it should be able to plan its path and help to tidy the kitchen, for example. So world model is a foundation that you can use to reason, to interact, and to create worlds.

中文翻译:

这只是其中的一部分，Lenny。我认为理解世界模型的一个简单方式是：这个模型允许任何人通过提示词（无论是图像还是句子），在脑海中创造出任何世界。并且能够在这个世界中互动，无论是在其中浏览行走、拿起物体还是改变事物，同时还能在这个世界中进行推理。例如，如果使用这个世界模型输出结果的主体是一个机器人，它应该能够规划路径并帮助清理厨房。所以，世界模型是一个你可以用来推理、互动和创造世界的“基础”。

[00:36:00] Lenny Rachitsky

English:

Great. Yeah. So robots feels like that's potentially the next big focus for AI researchers and just the impact on the world. And what you're saying here is this is a key missing piece of making robots actually work in the real world, understanding how the world works.

中文翻译:

太棒了。所以机器人似乎是AI研究者的下一个重大焦点，也将对世界产生巨大影响。而你在这里说的是，这是让机器人真正能在现实世界中工作的关键缺失环节——即理解世界是如何运作的。

[00:36:17] Dr. Fei-Fei Li

English:

Yeah. Well, first of all, I do think there's more than robots. That's exciting. But I agree with everything you just said. I think world modeling and spatial intelligence is a key missing piece of embodied AI. I also think let's not underestimate that humans are embodied agents and humans can be augmented by AI's intelligence. Just like today, humans are language animals, but we're very much augmented by AI helping us to do language tasks including software engineering. I think that we shouldn't underestimate or maybe we tend not to talk about how humans, as an embodied agents, can actually benefit so much from world models and spatial intelligence models as well as robots can.

中文翻译:

是的。首先，我认为除了机器人之外还有更多令人兴奋的应用。但我同意你刚才说的一切。我认为世界建模和空间智能是具身智能的关键缺失环节。同时，我们也不应低估人类本身也是“具身主体”，人类可以被AI的智能所增强。就像今天，人类是语言动物，而AI帮助我们处理语言任务（包括软件工程），极大地增强了我们的能力。我认为我们不应低估，或者说我们往往忽略了人类作为具身主体，其实也能像机器人一样，从世界模型和空间智能模型中获益良多。

[00:37:15] Lenny Rachitsky

English:

So the big unlocks here, robots, which a huge deal if this works out, imagine each of us has robots doing a bunch of stuff for us, they help us with disasters, things like that. Games obviously is a really cool example, just like infinitely playable games that you just invent out of your head. And then creativity feels like just like being fun, having fun, being creative, thinking of magic, wild new worlds, and environments.

中文翻译:

所以这里的重大突破点包括：机器人——如果成功了那将是大事，想象一下我们每个人都有机器人帮我们做事，帮我们处理灾难等等；游戏显然也是一个很酷的例子，就像你可以凭空想象出无限可玩的游戏；还有创造力，比如纯粹的趣味性、创造性，构思神奇、狂野的新世界和环境。

[00:37:39] Dr. Fei-Fei Li

English:

And also design, humans design from machines to buildings to homes and also scientific discovery. There is so much. I like to use the example of the discovery of the structure of DNA. If you look at one of the most important piece in DNA's discovery history is the x-ray diffraction photo that was captured by Rosalind Franklin, and it was a flat 2D photo of a structure that it looks like a cross with diffractions. You can google those photos. But with that 2D flat photo, the humans, especially two important humans, James Watson and Francis Crick, in addition to their other information, was able to reason in 3D space and deduce a highly three-dimensional double helix structure of the DNA. And that structure cannot possibly be 2D. You cannot think in 2D and deduce that structure. You have to think in 3D spatial, use the human spatial intelligence. So I think even in scientific discovery, spatial intelligence or AI-assisted spatial intelligence is critical.

中文翻译:

还有设计，人类设计从机器到建筑再到家居的一切，还有科学发现。应用场景太多了。我喜欢用 DNA 结构的发现作为例子。如果你看 DNA 发现史上最重要的证据之一，就是罗莎琳德·富兰克林 (Rosalind Franklin) 拍摄的那张 X 射线衍射照片。那是一张平面的 2D 照片，看起来像是一个带有衍射条纹的十字。你可以去搜一下。但凭借那张 2D 平面照片，人类——特别是詹姆斯·沃森 (James Watson) 和弗朗西斯·克里克 (Francis Crick) 这两位关键人物——结合其他信息，能够在 3D 空间中进行推理，并推导出 DNA 高度三维的双螺旋结构。那个结构不可能是 2D 的。你无法在 2D 思维中推导出那个结构，你必须在 3D 空间中思考，运用人类的空间智能。所以我认为即使在科学发现中，空间智能或 AI 辅助的空间智能也是至关重要的。

[00:39:08] Lenny Rachitsky

English:

This is such an example of, I think it was Chris Dixon that had this line that the next big thing is going to start off feeling like a toy. When ChatGPT just came out, I remember Sam Altman just tweeted it as like, "Here's a cool thing we're playing with, check it out." Now, it's the fastest growing product to all of history, changed the world. And it's oftentimes the things that just look like, okay, this is cool, that it's a fun to play with that end up changing the world most.

中文翻译:

这真是一个完美的例子。我想是 Chris Dixon 说过：“下一件大事刚开始时往往感觉像个玩具。”当 ChatGPT 刚出来时，我记得 Sam Altman 只是发推说：“这是我们正在玩的一个酷东西，大家来看看。”现在，它是历史上增长最快的产品，改变了世界。往往是那些看起来“挺酷、挺好玩”的东西，最终对世界的改变最大。

[00:39:33] Lenny Rachitsky (Sponsor: Sinch)

English:

This episode is brought to you by Sinch, the customer communications cloud. Here's the thing about digital customer communications. Whether you're sending marketing campaigns, verification codes, or account alerts, you need them to reach users reliably. That's where Sinch comes in. Over 150,000 businesses, including 8 of the top 10 largest tech companies globally use Sinch's API to build messaging, email, and calling into their products. And there's something big happening in messaging that product teams need to know about, Rich Communication Services or RCS. Think of RCS as SMS2.0. Instead of getting texts from a random number, your users will see your verified company name and logo without

needing to download anything new. It's a more secure and branded experience. Plus you get features like interactive carousels and suggested replies. And here's why this matters, US carriers are starting to adopt RCS. Sinch is already helping major brands send RCS messages around the world and they're helping Lenny's podcast listeners get registered first before the rush hits the US market. Learn more and get started at sinch.com/lenny. That's S-I-N-C-H.com/lenny.

中文翻译:

本期节目由客户沟通云平台 Sinch 赞助。关于数字化客户沟通，情况是这样的：无论你是发送营销活动、验证码还是账户提醒，你都需要它们可靠地送达用户。这就是 Sinch 的用武之地。超过 15 万家企业，包括全球前十大科技公司中的八家，都在使用 Sinch 的 API 将短信、邮件和通话功能集成到他们的产品中。现在，消息领域正在发生一件产品团队必须了解的大事：富通信服务 (RCS)。你可以把 RCS 看作是短信 2.0。用户收到的不再是来自随机号码的文本，而是能看到你经过验证的公司名称和 Logo，且无需下载任何新应用。这是一种更安全、更具品牌感的体验。此外，你还可以获得交互式轮播图和建议回复等功能。这之所以重要，是因为美国运营商正开始采用 RCS。Sinch 已经在帮助各大品牌在全球范围内发送 RCS 消息，他们正在帮助 Lenny 播客的听众在美国市场爆发前抢先注册。访问 sinch.com/lenny 了解更多并开始使用。

[00:40:45] Lenny Rachitsky

English:

I reached out to Ben Horowitz, who loves what you're doing, a big fan of yours. They're investors I believe in...

中文翻译:

我联系了 Ben Horowitz，他非常认可你正在做的事情，是你的超级粉丝。我相信他们投资了……

[00:40:51] Dr. Fei-Fei Li

English:

Yeah, we've known each other for many years, but yes, right now they're investors of World Labs.

中文翻译:

是的，我们认识很多年了，现在他们确实是 World Labs 的投资者。

[00:40:57] Lenny Rachitsky

English:

Amazing. Okay, so I asked him what I should ask you about and he suggested ask you why is the bitter lesson alone not likely to work for robots? So first of all, just explain what the bitter lesson was in the history of AI and then just why that won't get us to where we want to be with robots.

中文翻译:

太棒了。我问他我该问你什么，他建议我问你：为什么单靠“惨痛的教训”(The Bitter Lesson)可能对机器人行不通？所以首先，请解释一下 AI 历史上的“惨痛教训”是什么，然后解释为什么它无法带我们实现机器人的目标。

[00:41:17] Dr. Fei-Fei Li

English:

Well, first of all, there are many bitter lessons, but the bitter lessons everybody refers to is a paper written by Richard Sutton who won the Turing Award recently, and he does a lot of reinforcement learning. And Richard has said, if you look at the history, especially the algorithmic development of AI, it turns out simpler model with a ton of data always win at the end of the day instead of the more complex model with less data. I mean, that was actually... This paper came years after ImageNet. That to me was not bitter; it was a sweet lesson. That's why I built ImageNet because I believe that big data plays that role. So why can't bitter lesson work in robotics alone? Well, first of all, I think we need to give credit to where we are today. Robotics is very much in the early days of experimentation.

(00:42:25):

The research is not nearly as mature as say language models. So many people are still experimenting with different algorithms and some of those algorithms are driven by big data. So I do think big data will continue to play a role in robotics, but what is hard for robotics, there are a couple of things. One is that it's harder to get data. It's a lot harder to get data. You can say, well, there's web data. This is where the latest robotics research is using web videos. And I think web videos do play a role. But if you think about what made language model worth a very... As someone who does computer vision and spatial intelligence and robotics, I'm very jealous of my colleagues in language because they had this perfect setup where their training data are in words, eventually tokens, and then they produce a model that outputs words.

(00:43:36):

So you have this perfect alignment between what you hope to get, which we call objective function and what your training data looks like. But robotics is different. Even spatial intelligence is different. You hope to get actions out of robots, but your training data lacks actions in 3D worlds, and that's what robots have to do, right? Actions in 3D worlds. So you have to find different ways to fit a, what do they call, a square in a round hole, that what we have is tons of web videos. So then we have to start talking about adding supplementing data such as teleoperation data or synthetic data so that the robots are trained with this hypothesis of bitter lesson, which is large amount of data. I think there's still hope because even what we are doing in world modeling will really unlock a lot of this information for robots.

(00:44:53):

But I think we have to be careful because we're at the early days of this and bitter lesson is still to be tested because we haven't fully figured out the data for. Another part of the bitter lesson of robotics I think we should be so realistic about is again, compared to language models or even spatial models, robots are physical systems. So robots are closer to self-driving cars than a large language model. And that's very important to recognize. That means that in order for robots to work, we not only need brains, we also need the physical body. We also need application scenarios. If you look at the history of self-driving car, my colleague Sebastian Thrun took Stanford's car to win the first DARPA challenge in 2006 or 2005. It's 20 years since that prototype of a self-driving car being able to drive 130 miles in the Nevada desert to today's Waymo and on the street of San Francisco.

(00:46:17):

And we're not even done yet. There's still a lot. So that's a 20-year journey. And self-driving cars are much simpler robots, they're just metal boxes running on 2D surfaces, and the goal is not to touch anything. Robot is 3D things running in 3D world, and the goal is to touch things. So the journey is going to be, there's many aspects, elements, and of course one could say, well, the self-driving car, early algorithm were pre deep learning era. So deep learning is accelerating the brains. And I think that's true. That's why I'm in robotics, that's why I'm in spatial intelligence and I'm excited by it. But in the meantime, the car

industry is very mature and productizing also involves the mature use cases, supply chains, the hardware. So I think it's a very interesting time to work in these problems. But it's true, Ben is right. We might still be subject to a number of bitter lessons.

中文翻译:

首先，AI 历史上有很多教训，但大家通常指的“惨痛教训”是最近获得图灵奖的 Richard Sutton 写的一篇论文，他做了很多强化学习的研究。Richard 的观点是：回顾 AI 历史，特别是算法发展史，结果证明，拥有海量数据的简单模型最终总是能胜过数据较少的复杂模型。其实这篇论文是在 ImageNet 之后几年发表的。对我来说，这不算“惨痛”，而是一个“甜蜜的教训”。这就是我建立 ImageNet 的原因，因为我相信大数据的作用。那么，为什么“惨痛教训”不能直接解决机器人问题呢？首先，我们要正视现状：机器人技术目前仍处于实验的早期阶段。

(00:42:25):

它的研究远没有语言模型那么成熟。很多人还在尝试不同的算法，其中一些确实是由大数据驱动的。所以我认为大数据在机器人领域会继续发挥作用，但机器人面临的困难有几点。第一，获取数据更难，而且是难得多。你可以说有网页数据，最新的机器人研究确实在利用网页视频。我认为网页视频确实有用。但如果你想让语言模型如此成功……作为一个从事计算机视觉、空间智能和机器人研究的人，我非常嫉妒我的语言学同事，因为他们有一个完美的设定：训练数据是单词（最终是 Token），产出的模型输出的也是单词。

(00:43:36):

所以你的目标函数（你想得到的东西）和你的训练数据之间有着完美的对齐。但机器人不同，空间智能也不同。你希望机器人做出“动作”，但你的训练数据中缺乏 3D 世界中的动作数据，而这正是机器人必须做的：在 3D 世界中行动。所以你必须想办法把“方钉塞进圆孔”，因为我们拥有的是海量的网页视频。因此，我们必须开始讨论添加补充数据，比如远程操作数据（teleoperation data）或合成数据，以便按照“惨痛教训”的假设（即大量数据）来训练机器人。我认为仍有希望，因为我们正在做的世界建模将为机器人解锁大量此类信息。

(00:44:53):

但我觉得我们必须谨慎，因为我们还处于早期阶段，“惨痛教训”在机器人领域仍有待验证，因为我们还没完全解决数据来源问题。机器人领域另一个需要现实对待的“惨痛教训”是：与语言模型甚至空间模型相比，机器人是物理系统。机器人更接近自动驾驶汽车，而不是语言大模型。认识到这一点非常重要。这意味着为了让机器人工作，我们不仅需要大脑，还需要物理身体，还需要应用场景。看看自动驾驶的历史，我的同事 Sebastian Thrun 在 2005 或 2006 年带领斯坦福的赛车赢得了首届 DARPA 挑战赛。从那个能在内华达沙漠行驶 130 英里的自动驾驶原型，到今天的 Waymo 行驶在旧金山街头，已经过去了 20 年。

(00:46:17):

而且我们还没完全成功，还有很多路要走。那是 20 年的旅程。而自动驾驶汽车是相对简单的机器人，它们只是在 2D 平面上运行的金属盒子，目标是不碰到任何东西。而机器人是在 3D 世界中运行的 3D 物体，目标是去触碰东西。所以这段旅程会涉及很多方面和元素。当然，有人会说，自动驾驶早期的算法是在深度学习时代之前，而深度学习正在加速“大脑”的进化。我认为这是对的，这也是为什么我从事机器人和空间智能研究并为此感到兴奋。但与此同时，汽车行业非常成熟，产品化还涉及成熟的用例、供应链和硬件。所以我认为现在是研究这些问题的非常有趣的时期。但没错，Ben 是对的，我们可能仍会面临一系列“惨痛的教训”。

[00:47:28] Lenny Rachitsky

English:

Doing this work, do you ever just feel awe for the way the brain works and is able to do all of this for us? Just the complexity just to get a machine to just walk around and not hit things and fall, does just give you more respect for what we've already got?

中文翻译:

做这项工作时，你是否会对大脑的工作方式以及它为我们所做的一切感到敬畏？仅仅是让机器走动、不撞东西、不摔倒就如此复杂，这是否让你对我们天生拥有的能力更加尊重？

[00:47:44] Dr. Fei-Fei Li

English:

Totally. We operate on about 20 watts. That's dimmer than any light bulb in the room I'm in right now. And yet we can do so much. So I think actually the more I work in AI, the more I respect humans.

中文翻译:

完全正确。我们的运行功率大约只有 20 瓦，比我现在这个房间里任何一个灯泡都要暗。然而我们能做这么多事情。所以我觉得，我研究 AI 越多，我就越尊重人类。

[00:48:03] Lenny Rachitsky

English:

Let's talk about this product you just launched. It's called Marble, a very cute name. Talk about what this is, why this is important. I've been playing with it, it's incredible. We'll link to it for folks to check it out. What is Marble?

中文翻译:

让我们聊聊你刚刚发布的产品。它叫 Marble，一个很可爱的名字。聊聊这是什么，为什么它很重要。我一直在试用它，非常不可思议。我们会附上链接供大家查看。Marble 到底是什么？

[00:48:14] Dr. Fei-Fei Li

English:

Yeah, I'm very excited. So first of all, Marble is one of the first product that World Labs has rolled out. World Labs is a foundation frontier model company. We are founded by four co-founders who have deep technical history. My co-founders, Justin Johnson, Christoph Lassner, and Ben Mildenhall. We all come from the research field of AI, computer graphics, computer vision, and we believe that spatial intelligence and world modeling is as important, if not more, to language models and complementary to language models. So we wanted to seize this opportunity to create deep tech research lab that can connect the dots between frontier models with products. So Marble is an app that's built upon our frontier models. We've spent a year and plus building the world's first generative model that can output genuinely 3D worlds. That's a very, very hard problem.

(00:49:30):

And it was a very hard process and we have a team of incredible, founding team of incredible technologists from incredible teams. And then around just a month or two ago, we saw the first time that we can just prompt with a sentence and the image and multiple images and create worlds that we can just navigate in. If you put it on Google, which we have an option to let you do that, you can even walk around. Even though we've been building this for quite a while, it was still just awe-inspiring and we wanted to get into the hands of people who need it. And then we know that so many creators, designers, people who are thinking about robotic simulation, people who are thinking about different use cases of navigable interactive, immersive worlds game developers will find this useful. So we developed Marble

as a first step. It's again, still very early, but it's the world's first model doing this, and it's the world's first product that allows people to just prompt, we call it prompt to worlds.

中文翻译:

是的，我非常兴奋。首先，Marble 是 World Labs 推出的首批产品之一。World Labs 是一家基础前沿模型公司。我们由四位拥有深厚技术背景的创始人共同创立，我的合伙人包括 Justin Johnson、Christoph Lassner 和 Ben Mildenhall。我们都来自 AI、计算机图形学和计算机视觉研究领域。我们相信空间智能和世界建模与语言模型同样重要（甚至更重要），并且与语言模型互补。所以我们想抓住这个机会，建立一个深科技研究实验室，将前沿模型与产品连接起来。Marble 就是一个基于我们前沿模型的应用。我们花了一年多的时间构建了世界上第一个能够输出真正 3D 世界的生成式模型。这是一个非常、非常困难的问题。

(00:49:30):

这是一个非常艰难的过程，我们拥有一支由来自顶尖团队的优秀技术专家组成的创始团队。大约一两个月前，我们第一次看到只需输入一个句子、一张图片或多张图片，就能创造出我们可以直接在其中导航的世界。如果你把它放在 Google（我们提供这个选项），你甚至可以在里面走动。尽管我们已经开发了相当长一段时间，但看到那一幕依然令人惊叹，我们想把它交到需要它的人手中。我们知道很多创作者、设计师、从事机器人模拟的人、考虑可导航/可交互/沉浸式世界不同用例的人，以及游戏开发者都会发现它很有用。所以我们开发了 Marble 作为第一步。它目前还处于早期阶段，但它是世界上第一个做这件事的模型，也是第一个允许人们通过提示词生成世界（我们称之为 Prompt to Worlds）的产品。

[00:51:00] Lenny Rachitsky

English:

Well, I've been playing around with it. It is insane. You could just have a little Shire world where you just infinitely walk around middle earth basically, and there's no one there yet, but it's insane. You just go anywhere. There's dystopian world. I'm just looking at all these examples and my favorite part, actually, I don't know if there's a feature or bug, you can see the dots of the world before it actually renders with all the textures. And I just love like, you get a glimpse into what is going on with this model, basically-

中文翻译:

我一直在试玩，太疯狂了。你可以拥有一个小小的“夏尔”世界，基本上可以在中土世界无限走动，虽然里面还没人，但太震撼了。你可以去任何地方。还有反乌托邦世界。我正在看这些例子，我最喜欢的部分——其实我不知道这是功能还是 Bug——是在它完全渲染纹理之前，你可以看到构成世界的点阵。我非常喜欢这一点，因为你可以窥见这个模型内部到底在发生什么……

[00:51:27] Dr. Fei-Fei Li

English:

That is so cool to hear because this is where, as a researcher, I am learning because the dots that lead you into the world was an intentional feature visualization, is not part of the model. The model actually just generates the world. But we were trying to find a way to guide people into the world, and a number of engineers worked on different versions, but we converged on the dot, and so many people, you're not the only one, told us how delightful that experience is, and it was really satisfying for us to hear that this intentional visualization feature that's not just the big hardcore model actually has delighted our users.

中文翻译:

听到这个太酷了，因为作为研究者，我也在学习。那些引导你进入世界的点阵其实是一个刻意设计的“特征可视化”功能，它不是模型本身的一部分。模型实际上只是生成世界。但我们当时想找一种方式引导人们进入这

个世界，几位工程师尝试了不同版本，最后我们选定了点阵。很多人——你不是唯一一个——告诉我们这种体验多么令人愉悦。听到这个刻意设计的可视化功能（而不仅仅是那个硬核大模型本身）让用户感到快乐，我们非常满足。

[00:52:19] Lenny Rachitsky

English:

Wow. So you add that to make it more, like to have humans understand what's going on-

中文翻译:

哇。所以你们添加那个是为了让它更……比如让人们理解发生了什么……

[00:52:24] Dr. Fei-Fei Li

English:

To have fun, yes.

中文翻译:

是为了增加趣味性，是的。

[00:52:24] Lenny Rachitsky

English:

... get more delightful. Wow, that is hilarious. It makes me think about LLMs and the way they, it's not the same thing, but they talk about what they're thinking and what they're doing.

中文翻译:

……让它更讨喜。哇，太有意思了。这让我想起大语言模型（LLM），虽然不是一回事，但它们也会谈论自己在想什么、在做什么。

[00:52:32] Dr. Fei-Fei Li

English:

Yes, it is. It is.

中文翻译:

是的，确实如此。

[00:52:34] Lenny Rachitsky

English:

It also makes me think about just the Matrix. It's exactly the Matrix experience. I don't know if that was your inspiration.

中文翻译:

这也让我想到了《黑客帝国》(The Matrix)。这完全就是《黑客帝国》的体验。我不知道那是不是你们的灵感来源。

[00:52:42] Dr. Fei-Fei Li

English:

Well, like I said, a number of engineers worked on that. It could be their inspiration.

中文翻译:

就像我说的，好几位工程师参与了那个设计，那可能是他们的灵感。

[00:52:48] Lenny Rachitsky

English:

It's in their subconscious. Okay, so just for folks that may want to play around with this, maybe like, what are some applications today that folks can start using today? What's your goal with this launch?

中文翻译:

那是他们的潜意识。好，对于那些想尝试的人来说，目前有哪些应用是大家今天就可以开始使用的？你这次发布的目标是什么？

[00:52:59] Dr. Fei-Fei Li

English:

Yeah, so we do believe that world modeling is very horizontal, but we're already seeing some really exciting use cases, virtual production for movies, because what they need are 3D worlds that they can align with the camera. So when the actors are acting on it, they can position the camera and shoot the segments really well. And we're already seeing incredible use. In fact, I don't know if you have seen our launch video showing Marble. It was produced by a virtual production company. We collaborated with Sony and they use Marble scenes to shoot those videos. So we were collaborating with those technical artists and directors, and they were saying, this has cut our production time by 40X. In fact, it has to-

中文翻译:

是的，我们相信世界建模是非常通用的技术，但我们已经看到了一些非常令人兴奋的用例。比如电影的“虚拟制片”，因为他们需要能与摄像机对齐的3D世界。这样当演员表演时，他们可以定位摄像机并很好地拍摄片段。我们已经看到了惊人的应用。事实上，我不知道你是否看过我们展示Marble的发布视频，那是与一家虚拟制片公司合作制作的。我们与索尼合作，他们使用Marble场景来拍摄那些视频。我们与那些技术美术和导演合作，他们说这把制作时间缩短了40倍。事实上，它必须……

[00:53:00] Lenny Rachitsky

English:

40X?

中文翻译:

40倍？

[00:53:59] Dr. Fei-Fei Li

English:

Yes, in fact it has to, because we only had one month to work on this project and there were so many things they were trying to shoot. So using Marble really, really significantly accelerated the virtual production for VFX and movies. That's one use cases. We are already seeing our users taking our Marble scene and taking the mesh export and putting games, whether it's games on VR or just fun games that they have developed. We are showing an example of robotic simulation because when I was, I mean I still am a researcher doing robotic training. One of the biggest pain point is to create synthetic data for training robots. And this synthetic data needs to be very diverse. They need to come from different environments with different objects to manipulate. And one path to it is to ask computers to simulate.

(00:55:10):

Otherwise, humans have to build every single asset for robots. That's just going to take a lot longer. So we already have researchers reaching out and wanting to use Marble to create those synthetic environments. We also have unexpected user outreach in terms of how they want to use Marble. For example, a psychologist team called us to use Marble to do psychology research. It turned out some of the psychiatric patients they study, they need to understand how their brain respond to different immersive things of different features. For example, messy scenes or clean scenes or whatever you name it. And it's very hard for researchers to get their hands on these kind of immersive scenes and it will take them too long and too much budget to create. And Marble is a really almost instantaneous way of getting so many of these experimental environments into their hands. So we're seeing multiple use cases at this point. But the VFX, the game developers, the simulation developers as well as designers are very excited.

中文翻译:

是的，事实上必须如此，因为我们只有一个月的项目时间，而他们想拍的东西太多了。所以使用 Marble 显著加速了视觉特效 (VFX) 和电影的虚拟制片。这是一个用例。我们还看到用户导出 Marble 场景的网格 (mesh)，放入他们开发的 VR 游戏或趣味游戏中。我们还展示了一个机器人模拟的例子。作为一名从事机器人训练的研究者，最大的痛点之一就是为训练机器人创建“合成数据”。这些合成数据需要非常多样化，来自不同的环境，有不同的操作对象。一种途径就是让计算机去模拟。

(00:55:10):

否则，人类必须为机器构建每一个资产，那会耗费更长的时间。所以已经有研究人员联系我们，希望用 Marble 来创建这些合成环境。我们还收到了一些意想不到的用户反馈。例如，一个心理学家团队联系我们，想用 Marble 做心理学研究。事实证明，在他们研究的一些精神科病人中，他们需要了解大脑对具有不同特征的沉浸式场景（比如凌乱的场景或整洁的场景）如何反应。研究员很难获得这类沉浸式场景，自己制作又太耗时耗钱。而 Marble 提供了一种几乎瞬时的方式，将大量实验环境交到他们手中。所以目前我们看到了多种用例，视觉特效、游戏开发、模拟开发以及设计师们都非常兴奋。

[00:56:39] Lenny Rachitsky

English:

This is very much the way things work in AI. I've had other AI leaders on the podcast and it's always put things out there early as soon as you can to discover where the big use cases are. The head of ChatGPT told me how, when they first put out ChatGPT, he was just scanning TikTok to see how people were using it and all the things they were talking about, and that's what convinced them where to lean in and help them see how people actually want to use it. I love this last use case for therapy. I'm just imagining heights, people dealing with heights or snakes or spiders, which-

中文翻译:

这非常符合 AI 领域的工作方式。我请过其他 AI 领袖上播客，他们总是说要尽早发布产品，以便发现重大的用例。ChatGPT 的负责人告诉我，当他们刚发布 ChatGPT 时，他一直在刷 TikTok，看人们怎么用它，看大家在聊什么，正是这些让他确定了该往哪里发力，并看清了人们真正的需求。我非常喜欢那个治疗的用例。我能想象恐高症患者，或者怕蛇、怕蜘蛛的人……

[00:57:11] Dr. Fei-Fei Li

English:

It's amazing. A friend of mine last night literally called me and talked about his height scare and asked me if Marble should be used. It's amazing you went straight there.

中文翻译:

太神奇了。昨晚我一个朋友真的给我打电话，聊到他恐高，问我能不能用 Marble。你竟然直接想到了这一点，太神了。

[00:57:24] Lenny Rachitsky

English:

Because imagining all the exposure therapy stuff, this could be so good for that. That is so cool. Okay, so I should have asked you this before, but I think there's going to be a question of just, how does this differ from things like VO3 and other video generation models? It's pretty clear to me, but I think it might be helpful just to explain how this is different from all the video AI tools people have seen.

中文翻译:

因为想到那些“暴露疗法”(exposure therapy)，这个工具会非常合适。太酷了。好，我之前应该问你的，但我认为大家会有一个疑问：这与 Sora 或其他视频生成模型有什么不同？对我来说区别很明显，但我觉得解释一下它与人们见过的视频 AI 工具有何不同会很有帮助。

[00:57:46] Dr. Fei-Fei Li

English:

World Labs' thesis is that spatial intelligence is fundamentally very important, and spatial intelligence is not just about videos. In fact, the world is not passively watching videos passing by. I love, Plato has the allegory of the cave analogy to describe vision. He said that imagine a prisoner tied on his chair, not very humane, but in a cave watching a full life theater in front of him, but the actual life theater that actors are acting is behind his back. It was just lit so that the projection of the action is on a wall of the cave. And then the goal, the task of this prisoner is to figure out what's going on. It's a pretty extreme example, but it really shows, it describes what vision is about, is that to make sense of the 3D world or 4D world out of 2D. So spatial intelligence to me is deeper than only creating that flat 2D world.

(00:59:14):

Spatial intelligence to me is the ability to create, reason, interact, make sense of deeply spatial world, whether it's 2D or 3D or 4D, including dynamics and all that. So World Lab is focusing on that, and of course the ability to create videos per se could be part of this. And in fact, just a couple of weeks ago, we rolled out the world's first real time demoable, real-time video generation on a single H100 GPU. So part of our technology includes that, but I think Marble is very different because we really want creators,

designers, developers to have in their hands a model that can give them worlds with 3D structures so they can use it for their work. And that's why Marble is so different.

中文翻译:

World Labs 的核心论点是：空间智能从根本上非常重要，而空间智能不仅仅关乎视频。事实上，世界并不是在被动地观看流逝的视频。我非常喜欢柏拉图用“洞穴寓言”来描述视觉。他说，想象一个被绑在椅子上的囚犯（虽然不太人道），他在洞穴里看着面前的一场生活剧，但演员们表演的真实场景其实在他身后，只是被火光照耀，将动作投射到了洞穴的墙壁上。这个囚犯的任务就是弄清楚发生了什么。这是一个极端的例子，但它真实地描述了视觉的本质：从 2D 投影中理解 3D 或 4D 世界。所以对我来说，空间智能比仅仅创建平面的 2D 世界要深刻得多。

(00:59:14):

空间智能对我来说是创造、推理、互动和理解深度空间世界的能力，无论它是 2D、3D 还是 4D（包括动态等）。World Labs 专注于此。当然，生成视频本身可以是其中的一部分。事实上，就在几周前，我们推出了世界上第一个可以在单块 H100 GPU 上实时演示的实时视频生成技术。所以我们的技术包含这一部分，但我认为 Marble 非常不同，因为我们真心希望创作者、设计师、开发者手中能拥有一个可以提供 3D 结构世界的模型，以便他们用于工作。这就是 Marble 如此独特的原因。

[01:00:21] Lenny Rachitsky

English:

The way I see it is it's a platform for a ton of opportunity to do stuff. As you described, videos are just like, here's a one-off video that's very fun and cool and you could... And that's it. That's it. And you move on.

中文翻译:

我的理解是，它是一个提供大量机会去做事的平台。正如你所描述的，视频就像是“这是一个一次性的视频，很有趣很酷”，然后就没了，你接着看下一个。

[01:00:33] Dr. Fei-Fei Li

English:

By the way, we could in Marble, we could allow people to export in video forms. So you could actually, like you said, you go into a world, so let's say it's a hobbit cave. You can actually, especially as a creator, you have such a specific way of moving the camera in a trajectory in the director's mind, and then you can export that from Marble into a video.

中文翻译:

顺便说一下，在 Marble 中，我们也允许人们以视频形式导出。所以你可以进入一个世界，比如一个霍比特人洞穴，作为创作者，你可以按照导演脑海中的特定轨迹移动摄像机，然后将这段过程从 Marble 导出为视频。

[01:01:02] Lenny Rachitsky

English:

What does it take to create something like this? Just how big is the team, how many GPUs you work in? Anything you can share there. I don't know how much of this is private information, but just what does it take to create something like this that you've launched here?

中文翻译:

创造这样的东西需要什么？团队规模有多大？用了多少 GPU？有什么可以分享的吗？我不知道这其中有多少是机密，但创造这样一个产品需要投入什么？

[01:01:12] Dr. Fei-Fei Li

English:

It takes a lot of brain power. So we just talk about 20 watts per brain. So from that point of view, it's a small number, but it's actually incredible. It's half billion years of evolution to give us those power. We have a team of 30-ish people now, and we are predominantly researchers and research engineers, but we also have designers and product. We actually really believe that we want to create a company that's anchored in the deep tech of spatial intelligence, but we are actually building serious products. So we have this integration of R&D and productization, and of course, we use a ton of GPUs.

中文翻译:

这需要大量的脑力。我们刚才聊到每个大脑只有 20 瓦。从这个角度看，数字很小，但其实很不可思议，那是经过 5 亿年进化才赋予我们的力量。我们现在有一个大约 30 人的团队，主要是研究员和研究工程师，但也有设计师和产品人员。我们坚信要建立一家立足于空间智能深层技术、同时构建严肃产品的公司。所以我们将研发与产品化结合在一起。当然，我们也使用了大量的 GPU。

[01:02:15] Lenny Rachitsky

English:

That's the technical thing.

中文翻译:

那是技术层面的。

[01:02:17] Dr. Fei-Fei Li

English:

Happy to hear.

中文翻译:

很高兴听你这么说。

[01:02:20] Lenny Rachitsky

English:

Well, congrats on the launch. I know this is a huge milestone. I know this took a ton of work.

中文翻译:

祝贺产品发布。我知道这是一个巨大的里程碑，肯定投入了大量心血。

[01:02:20] Dr. Fei-Fei Li

English:

Thank you.

中文翻译:

谢谢。

[01:02:23] Lenny Rachitsky

English:

So I just want to say congrats to you and your team. Let me talk about your founder journey for a moment. So you're a founder of this company. You started how many years ago? A couple of years ago, two, three years ago?

中文翻译:

再次祝贺你和你的团队。让我们聊聊你的创业历程。你是这家公司的创始人，你是几年前开始的？两三年前？

[01:02:23] Dr. Fei-Fei Li

English:

A year ago.

中文翻译:

一年前。

[01:02:33] Lenny Rachitsky

English:

A year ago?

中文翻译:

一年前？

[01:02:34] Dr. Fei-Fei Li

English:

A year plus.

中文翻译:

一年多。

[01:02:37] Lenny Rachitsky

English:

A year? Okay. Wow.

中文翻译:

一年？好的，哇。

[01:02:37] Dr. Fei-Fei Li

English:

Probably, 18 month, yeah.

中文翻译:

大概 18 个月吧，是的。

[01:02:38] Lenny Rachitsky

English:

Okay. What's something you wish you knew before you started this that you wish you could whisper into the ear of Fei-Fei of 18 months ago?

中文翻译:

好。有什么是你希望在开始之前就知道的，或者说你希望对 18 个月前的飞飞耳语些什么？

[01:02:46] Dr. Fei-Fei Li

English:

Well, I continue to wish I know the future of technology. I think actually that's one of our founding advantage is that we see the future earlier in general than most people. But still, man, this is so exciting and so amazing that what's unknown and what's coming, but I know the reason you're asking me this question is not about the future of technology. Furthermore, look, I did not start a company of this scale at 20-year-old. So I started a dry cleaner when I was 19, but that's a little smaller scale.

中文翻译:

嗯，我一直希望自己能预知技术的未来。我认为我们创业的优势之一就是通常比大多数人更早看到未来。但即便如此，天呐，那些未知的、即将到来的事物依然如此令人兴奋和惊叹。但我知道你问这个问题不是为了聊技术未来。此外，你看，我并不是在 20 岁时创办这种规模的公司的。我 19 岁时开过一家洗衣店，但那规模要小得多。

[01:03:30] Lenny Rachitsky

English:

We got to talk about that.

中文翻译:

我们得聊聊那个。

[01:03:32] Dr. Fei-Fei Li

English:

And then I founded Google Cloud AI and then I founded an institute at Stanford but those are different beasts. I did feel I was a little more prepared as a founder of the grinding journey compared to maybe the

20-year-old founders. But I still, I'm surprised, and it puts me into paranoia sometimes that how intensely competitive AI landscape is from the model, the technology itself, as well as talents. And when I founded the company, we did not have these incredible stories of how much certain talents would cost. So these are things that continue to surprise me and I have to be very alert about.

中文翻译:

后来我创办了 Google Cloud AI，然后在斯坦福创办了一个研究院，但那些都是不同的性质。与 20 岁的创始人相比，我觉得自己作为创始人对这段艰辛旅程的准备更充分一些。但我依然感到惊讶，有时甚至会感到一种危机感 (paranoia)，因为 AI 领域的竞争太激烈了——从模型、技术本身到人才。当我创办公司时，还没听说过某些人才的身价会高到如此离谱。这些事情一直让我感到惊讶，我必须保持高度警惕。

[01:04:40] Lenny Rachitsky

English:

So the competition you're talking about is the competition for talent, the speed at which just how things are moving.

中文翻译:

所以你说的竞争是指人才竞争，以及事物发展的速度。

[01:04:46] Dr. Fei-Fei Li

English:

Yeah.

中文翻译:

是的。

[01:04:47] Lenny Rachitsky

English:

Yeah. You mentioned this point that I want to come back to that if you just look over the course of your career, you were at all of the major collections of humans that led to so many of the breakthroughs that are happening today. Obviously, we talk about ImageNet also just SAIL at Stanford is where a lot of the work happened, Google Cloud, which a lot of the breakthroughs happened. What brought you to those places? Like for people looking for how to advance in their career, be at the center of the future, just is there a through line there of just what pulled you from place to place and pulled you into those groups that might be helpful for people to hear?

中文翻译:

是的。你提到了一点我想再聊聊：纵观你的职业生涯，你曾身处所有导致今天重大突破的关键人才聚集地。显然，我们谈到了 ImageNet，还有斯坦福的 SAIL 实验室（很多工作在那里完成），以及 Google Cloud（诞生了许多突破）。是什么把你带到这些地方的？对于那些想要提升职业生涯、处于未来核心的人来说，有没有一条主线，或者说是什么吸引你从一个地方到另一个地方，并进入那些团队的？这对听众可能会很有帮助。

[01:05:25] Dr. Fei-Fei Li

English:

Yeah, this is actually a great question, Lenny, because I do think about it, and obviously we talked about it's curiosity and passion that brought me to AI, that is more a scientific north star, right? I did not care if AI was a thing or not, so that was one part. But how did I end up choosing in the particular places I work in, including starting World Labs, is I think I'm very grateful to myself or maybe to my parents' genes. I'm an intellectually very fearless person, and I have to say when I hire young people, I look for that because I think that's a very important quality if one wants to make a difference, is that when you want to make a difference, you have to accept that you're creating something new or you're diving into something new. People haven't done that. And if you have that self-awareness, you almost have to allow yourself to be fearless and to be courageous.

(01:06:42):

So when I, for example, came to Stanford, in the world of academia, I was very close to this thing called tenure, which is have the job forever at Princeton. But I chose to come to Stanford because... I love Princeton. It's by alma mater. It's just at that moment there are people who are so amazing at Stanford and the Silicon Valley ecosystem was so amazing that I was okay to take a risk of restarting my tenure clock. Becoming the first female director of SAIL, I was actually relatively speaking a very young faculty at that time, and I wanted to do that because I care about that community. I didn't spend too much time thinking about all the failure cases.

(01:07:46):

Obviously, I was very lucky that the more senior faculty supported me, but I just wanted to make a difference. And then going to Google was similar. I wanted to work with people like Jeff Dean, Jeff Hinton, and all these incredible demists, the incredible people. The same with World Labs. I have this passion. And I also believe that people with the same mission can do incredible things. So that's how it guided my through line. I don't overthink of all possible things that can go wrong because that's too many.

中文翻译:

这确实是个好问题，Lenny。我确实思考过。显然，我们谈过是好奇心和热情把我带到了 AI 领域，那是科学上的北极星，对吧？我当时并不在乎 AI 是否会成为热门。但我是如何选择具体工作地点的，包括创办 World Labs？我想我非常感激自己，或者说感激父母的基因。我在智力上是一个非常无畏的人。我必须说，当我招聘年轻人时，我会寻找这种特质，因为我认为如果你想有所作为，这是一个非常重要的品质。当你想要有所作为时，你必须接受你正在创造新事物或投身于新事物，而那是别人没做过的。如果你有这种自觉，你几乎必须允许自己变得无畏和勇敢。

(01:06:42):

例如，当我来到斯坦福时，在学术界，我当时在普林斯顿大学已经快拿到终身教职（Tenure）了，那意味着一份永久的工作。但我选择来到斯坦福，因为……我爱普林斯顿，那是我的母校。只是在那个时刻，斯坦福的人才太出色了，硅谷的生态系统太迷人了，我愿意冒着重新开始终身教职评定周期的风险。成为 SAIL 的第一位女性主任时，相对而言我当时还是一名非常年轻的教员，但我愿意去做，因为我关心那个社区。我没有花太多时间去想所有失败的可能性。

(01:07:46):

显然，我很幸运得到了资深教员的支持，但我当时只是想有所作为。去 Google 也是类似的，我想和 Jeff Dean、Geoff Hinton 以及所有那些了不起的人一起工作。创办 World Labs 也是一样。我有这种热情，我也相信志同道合的人可以成就伟业。这就是引导我的主线。我不会过度思考所有可能出错的事情，因为那实在太多了。

[01:08:33] Lenny Rachitsky

English:

I feel like an important element of this is not focusing on the downside, focusing more on the people, the mission. What gets you excited, what do you think, the curiosity.

中文翻译:

我觉得其中一个重要因素是不去关注负面影响，而是更多地关注人、关注使命。关注什么让你兴奋，关注你的好奇心。

[01:08:43] Dr. Fei-Fei Li

English:

Yeah. I do want to say one thing to all the young talents in AI, the engineers, the researchers out there, because some of you apply to World Labs, I feel very privileged you considered World Labs. I do find many of the young people today think about every single aspect of an equation when they decide on jobs. At some point, maybe that's the way they want to do it, but sometimes I do want to encourage young people to focus on what's important because I find myself constantly in mentoring mode when I talk to job candidates. Not necessarily recruiting or not recruiting, but just in mentoring mode when I see an incredible young talent who is over-focusing on every minute dimension and aspect of considering a job, when maybe the most important thing is where's your passion? Do you align with the mission? Do you believe and have faith in this team? And just focus on the impact and you can make and the kind of work and team you can work with.

中文翻译:

是的。我想对所有 AI 领域的年轻人才、工程师和研究人员说一句话，因为你们中有些人申请了 World Labs，我很荣幸你们能考虑我们。我发现现在的很多年轻人在决定工作时，会考虑方程中的每一个细微变量。在某种程度上，也许那是他们的方式，但有时我想鼓励年轻人专注于重要的事情。当我与求职者交谈时，我发现自己经常处于“导师模式”。不一定是为了招聘，只是当我看到一个优秀的年轻人才过度关注工作的每一个细枝末节时，我会想：也许最重要的事情是你的热情在哪里？你是否认同这个使命？你是否相信并对这个团队有信心？专注于你能产生的影响，以及你能从事的工作和团队。

[01:10:05] Lenny Rachitsky

English:

Yeah, it's tough. It's tough for people in the AI space. Now there's so much, so much at them, so much new, so much happening, so much FOMO.

中文翻译:

是的，这很难。对于 AI 领域的人来说很难。现在有太多的信息冲击着他们，太多的新事物，太多的变化，还有太多的 FOMO（错失恐惧症）。

[01:10:11] Dr. Fei-Fei Li

English:

That's true.

中文翻译:

确实如此。

[01:10:12] Lenny Rachitsky

English:

I could see the stress. And so I think that advice is really important. Just like what will actually make you feel fulfilled in what you're doing, not just where's the fastest growing company, where's the... Who's going to win? I don't know. I want to make sure I ask you about the work you're doing today at Stanford, at the HCI. I think it's the-

中文翻译:

我能感受到那种压力。所以我认为那个建议非常重要：什么能让你在工作中感到充实，而不仅仅是哪家公司增长最快，或者谁会赢。我不知道。我还想问问你在斯坦福 HAI 所做的工作。我想那是……

[01:10:12] Dr. Fei-Fei Li

English:

HAI.

中文翻译:

是 HAI。

[01:10:30] Lenny Rachitsky

English:

HAI, Human-Centered AI Institute. What are you doing there? I know this is a thing you do on the side still.

中文翻译:

HAI，人文 AI 研究院。你在那里做什么？我知道你现在依然在兼顾那里的工作。

[01:10:36] Dr. Fei-Fei Li

English:

So yes, HAI, Human-Centered AI Institute was co-founded by me and a group of faculty like Professor John Etchemendy, Professor James Landay, Professor Chris Manning back in 2018. I was actually finishing my last sabbatical at Google and it was a very, very important decision for me because I could have stayed in industry, but my time at Google taught me one thing is AI is going to be a civilization of technology. And it dawned on me how important this is to humanity to the point that I actually wrote a piece in New York Times, that year 2018, to talk about the need for a guiding framework to develop and to apply AI. And that framework has to be anchored in human benevolence, in human centeredness. And I felt that Stanford, one of the world's top university in the heart of Silicon Valley that gave birth to important companies from NVIDIA to Google, should be a thought leader to create this human-centered AI framework and to actually embody that in our research education and policy and ecosystem work.

(01:12:10):

So I founded HAI. Fast-forward, after six, seven years, it has become the world's largest AI institute that does human-centered research, education, ecosystem, outreach, and policy impact. It involves hundreds of faculty across all eight schools at Stanford, from medicine to education, to sustainability to business, to engineering, to humanities to law. And we support researchers, especially at the interdisciplinary area from digital economy, to legal studies, to political science, to discovery of new drugs, to new algorithms to that's beyond transformers. We also actually put a very strong focus on policy because when we started HAI, I realized that Silicon Valley did not talk to Washington DC and or Brussels or other parts of the world.

(01:13:27):

And given how important this technology is, we need to bring everybody on board. So we created multiple programs from congressional bootcamp to AI index report to policy briefing, and we especially participated in policymaking including advocating for a national AI research cloud bill that was passed in the first Trump administration and participating in state level regulatory AI discussions. So there's a lot we did, and I continue to be one of the leaders even though I'm much less involved operationally because I care not only we create this technology, but we use it in the right way.

中文翻译:

是的，斯坦福人文 AI 研究院（HAI）是由我和约翰·埃切门迪（John Etchemendy）教授、詹姆斯·兰迪（James Landay）教授、克里斯·曼宁（Chris Manning）教授等一批教员在 2018 年共同创立的。当时我正结束在 Google 的学术休假，那对我来说是一个非常重要的决定，因为我本可以留在工业界。但在 Google 的经历教会了我一件事：AI 将成为一项文明级别的技术。我意识到这对人类是多么重要，以至于我在 2018 年给《纽约时报》写了一篇文章，讨论在开发和应用 AI 时需要一个指导框架。这个框架必须植根于人类的善意和以人为本。我觉得斯坦福作为位于硅谷心脏地带的世界顶尖大学，孕育了从 NVIDIA 到 Google 的重要公司，理应成为思想领袖，创建这个以人为本的 AI 框架，并将其体现在我们的研究、教育、政策和生态工作中。

(01:12:10):

于是我创立了 HAI。快进到六七年后的今天，它已成为全球最大的从事以人为本 AI 研究、教育、生态拓展和政策影响的机构。它涉及斯坦福所有八个学院的数百名教员，从医学到教育，从可持续发展到商业，从工程到人文再到法律。我们支持跨学科领域的研究人员，从数字经济到法律研究，从政治科学到新药研发，再到超越 Transformer 的新算法。我们还非常重视政策，因为在创办 HAI 时，我意识到硅谷并不怎么与华盛顿特区、布鲁塞尔或世界其他地方沟通。

(01:13:27):

考虑到这项技术的重要性，我们需要让所有人参与进来。因此，我们创建了多个项目，从国会训练营到 AI 指数报告，再到政策简报。我们特别参与了政策制定，包括倡导在第一届特朗普政府期间通过的国家 AI 研究云法案（NAIRR），以及参与州一级的 AI 监管讨论。我们做了很多工作，虽然我现在较少参与具体运营，但我依然是领导者之一，因为我不仅关心我们创造了这项技术，更关心我们以正确的方式使用它。

[01:14:24] Lenny Rachitsky

English:

Wow. I was not aware of all that other work you were doing. As you're talking, I was reminded Charlie Munger had this quote, "Take a simple idea and take it very seriously." I feel like you've done that in so many different ways and stayed with it and it's unbelievable the impact that you've had in so many ways over the years. I'm going to skip the lightning round and I'm just looking to ask you one last question. Is there anything else that you wanted to share? Anything else you want to leave listeners with?

中文翻译:

哇。我以前不知道你还做了这么多其他工作。听你说话时，我想起了查理·芒格的一句话：“找一个简单的想法，并极其严肃地对待它。”我觉得你在很多不同的方面都做到了这一点，并且坚持了下来。这些年来你在这么多领域产生的影响力简直不可思议。我打算跳过闪电问答环节，直接问最后一个问题是：你还有什么想分享的吗？有什么想留给听众的话吗？

[01:14:52] Dr. Fei-Fei Li

English:

I am very excited by AI, Lenny. I want to answer one question that when I travel around the world, everybody asks me is that, if I'm a musician, if I'm a teacher, middle school teacher, if I'm a nurse, if I'm an accountant, if I'm a farmer, do I have a role in AI or is AI just going to take over my life or my work? And I think this is the most important question of AI and I find that in Silicon Valley, we tend not to speak heart-to-heart with people, with people like us and not like us in Silicon Valley, but all of us, we tend to just toss around words like infinite productivity or infinite leisure time or infinite power or whatever. But at the end of the day, AI is about people. And when people ask me that question, it's a resounding yes, everybody has a role in AI.

(01:16:03):

It depends on what you do and what you want. But no technology should take away human dignity and the human dignity and agency should be at the heart of the development, the deployment, as well as the governance of every technology. So if you are a young artist and your passion is storytelling, embrace AI as a tool. In fact, embrace Marble. I hope it becomes a tool for you because the way you tell your story is unique and the world still needs it. But how you tell your story, how do you use the most incredible tool to tell your story in the most unique way is important. And that voice needs to be heard. If you are a farmer near retirement, AI still matters because you are a citizen. You can participate in your community, you should have a voice in how AI is used, how AI is applied.

(01:17:14):

You work with people that you can encourage all of you to use AI to make life easier for you. If you are a nurse, I hope you know that at least in my career, I have worked so much in healthcare research because I feel our healthcare workers should be greatly augmented and helped by AI technology, whether it's smart cameras to feed more information or robotic assistance because our nurses are overworked, overfatigued, and as our society ages, we need more help for people to be taken care of. So AI can play that role. So I just want to say that it's so important that even a technologist like me are sincere about that everybody has a role in AI.

中文翻译:

Lenny，我对AI感到非常兴奋。我想回答一个我环游世界时每个人都会问我的问题：如果我是一名音乐家、一名中学老师、一名护士、一名会计或一名农民，我在AI时代还有一席之地吗？还是说AI只会接管我的生活或工作？我认为这是AI最重要的问题。我发现，在硅谷，我们往往不习惯与人交心——无论是与硅谷内部还是外部的人。我们倾向于随口抛出“无限生产力”、“无限闲暇时间”或“无限权力”之类的词。但归根结底，AI关乎于人。当人们问我那个问题时，我的回答是肯定的：每个人在AI时代都有自己的角色。

(01:16:03):

这取决于你做什么以及你想要什么。但任何技术都不应剥夺人的尊严，人的尊严和自主权（agency）应该处于每项技术开发、部署和治理的核心。所以，如果你是一个热爱讲故事的年轻艺术家，请拥抱AI这个工具。事实上，去拥抱Marble吧，我希望它能成为你的工具，因为你讲故事的方式是独一无二的，世界依然需要它。但你如何讲故事，如何利用最不可思议的工具以最独特的方式讲故事，这很重要。那个声音需要被听到。如果你是

一个快退休的农民，AI 依然重要，因为你是一个公民。你可以参与社区，你应该在 AI 如何使用、如何应用方面拥有发言权。

(01:17:14):

你可以鼓励大家使用 AI 来让生活变得更轻松。如果你是一名护士，我希望你知道，至少在我的职业生涯中，我投入了大量精力在医疗保健研究上，因为我觉得我们的医护人员应该得到 AI 技术的极大增强和帮助。无论是通过智能摄像头提供更多信息，还是通过机器人辅助，因为我们的护士工作过度、极度疲劳，而随着社会老龄化，我们需要更多帮助来照顾人们。AI 可以扮演这个角色。所以我想说，即使像我这样的技术专家，也非常真诚地认为：每个人在 AI 时代都有自己的角色。

[01:18:17] Lenny Rachitsky

English:

What a beautiful way to end it. Such a tie back to where we started about how it's up to us and take individual responsibility for what AI will do in our lives. Final question, where can folks find Marble? Where can they go, maybe try to join World Labs if they want to? What's the website? Where do people go?

中文翻译:

多么完美的结尾。这正好呼应了我们开始时聊到的：这取决于我们，我们要为 AI 在我们生活中的作用承担个人责任。最后一个问题，大家在哪里可以找到 Marble？如果有人想加入 World Labs，该去哪里？网站是什么？

[01:18:34] Dr. Fei-Fei Li

English:

Well, World Labs website is www.worldlabs.ai and you can find our research progress there. We have technical blogs. You can find Marble, the product there. You can sign in there. You can find our job posts link there. We're in San Francisco. We love to work with the world's best talents.

中文翻译:

World Labs 的网站是 www.worldlabs.ai。你可以在那里看到我们的研究进展，我们有技术博客。你可以在那里找到 Marble 这个产品并登录体验。你也可以在那里找到招聘链接。我们在旧金山，我们非常渴望与世界上最优秀的人才合作。

[01:19:02] Lenny Rachitsky

English:

Amazing. Fei-Fei, thank you so much for being here.

中文翻译:

太棒了。飞飞，非常感谢你能来。

[01:19:04] Dr. Fei-Fei Li

English:

Thank you, Lenny.

中文翻译:

谢谢你, Lenny。

[01:19:06] Lenny Rachitsky

English:

Bye everyone.

(01:19:09):

Thank you so much for listening. If you found this valuable, you can subscribe to the show on Apple Podcasts, Spotify, or your favorite podcast app. Also, please consider giving us a rating or leaving a review as that really helps other listeners find the podcast. You can find all past episodes or learn more about the show at lennyspodcast.com. See you in the next episode.

中文翻译:

大家再见。

(01:19:09):

非常感谢收听。如果你觉得本期节目有价值, 可以在 Apple Podcasts、Spotify 或你喜欢的播客应用中订阅。此外, 请考虑给我们评分或留下评论, 这能极大地帮助其他听众发现这个播客。你可以在 lennyspodcast.com 找到所有往期节目或了解更多信息。下期节目再见。