

RAMESH JOHARI

LENNY'S PODCAST

BILINGUAL TRANSCRIPT

ORIGINAL BY

Lenny Rachitsky

@lennysan • x.com/lennysan

ANALYSIS BY

@Penny777 • x.com/penny777

Ramesh Johari - 双语对照

This is the complete bilingual transcript for Lenny's Podcast featuring Ramesh Johari.

[00:00:00] [Ramesh Johari]

English:

Marketplaces are a little bit like a game of whac-a-mole. One example that I came across with one of the companies I worked with that I love is our new supply side was having a pretty bad experience. So what we decided to do is build some custom bespoke features that were really going to direct them to more experienced folks on the other side of the market. Good. And then, yeah, lo and behold, pretty soon those metrics start to look better. But then we're looking at it, we're like, "Wait a second. Now the existing folks on the other side are having a worse experience," so you kind of whiplash around. You're like, "Oh, wait a second. We better do something about that." So we take them, we try to match them up with the more experienced folks, and now suddenly a month after that, you're like, "Wait a second," and your metrics just keep moving around. And that's because the whac-a-mole game here is ultimately, a lot of marketplace management is moving attention and inventory around. Many of the changes that are most consequential create winners and losers. And rolling with those changes is about recognizing whether the winners you've created are more important to your business than the losers you've created in the process.

中文翻译:

双边市场（Marketplaces）有点像打地鼠游戏。我合作过的一家公司有一个我很喜欢的例子：当时我们的新供应端（Supply side）体验非常糟糕。于是我们决定开发一些定制化的功能，将他们引导至市场另一端更有经验的用户。结果，你瞧，很快这些指标就开始好转了。但接着我们观察发现，“等一下，现在另一端原有的资深用户体验变差了。”于是你就像遭遇了挥鞭效应（Whiplash）一样，心想：“噢，等等，我们得处理一下。”于是我们又去调整，尝试把他们与更有经验的人匹配。结果一个月后，你又会惊呼：“等等，”你的指标就这样不停地变动。这是因为打地鼠游戏的本质在于，双边市场的管理最终其实是在调配注意力和库存。许多影响深远的变革都会产生赢家和输家。而应对这些变化的关键在于，你要意识到在这个过程中，你创造的赢家是否比你产生的输家对你的业务更重要。

[00:01:00] [Lenny]

English:

Today my guest is Ramesh Johari. Ramesh is a professor at Stanford University, where he does research on and teaches data science methods and practices with a specific focus on the design and operation of online marketplaces. He's advised and worked with some of the biggest marketplaces in the world, including Airbnb, Uber, Stripe, Bumble, Stitch Fix, Upwork, and many others. And in our conversation, we get super nerdy on how to build a thriving marketplace, including where to focus your resources to fuel the marketplace flywheel growth, why data and data science is so central to building a successful marketplace, how to design a better review system. Why as a founder, you shouldn't think of yourself as a marketplace founder, but instead simply as a founder. Also, how AI is going to impact data science and

marketplaces, and experimentation, and so much more. If you're building a marketplace business, or thinking about building a marketplace, or just curious, this episode is for you. With that, I bring you Ramesh Johari after a short word from our sponsors.

中文翻译:

今天的嘉宾是 Ramesh Johari。Ramesh 是斯坦福大学的教授，他在那里研究并教授数据科学的方法与实践，特别关注在线双边市场的设计与运营。他曾为全球一些最大的平台提供咨询或合作，包括 Airbnb、Uber、Stripe、Bumble、Stitch Fix、Upwork 等等。在我们的对话中，我们将深入探讨如何构建一个繁荣的市场，包括：将资源集中在哪里以驱动市场飞轮增长；为什么数据和数据科学对构建成功市场如此核心；如何设计更好的评价系统；为什么作为创始人，你不应该把自己定位为“市场型创始人”，而仅仅是一个“创始人”。此外，我们还会聊到 AI 将如何影响数据科学、市场运营和实验等等。如果你正在构建双边市场业务，或者正考虑这么做，亦或是纯粹好奇，这一集就是为你准备的。在听完赞助商的简短介绍后，让我们欢迎 Ramesh Johari。

[00:02:04] [Lenny]

English:

This episode is brought to you by Sanity. Your website is the heart of your growth engine. For that engine to drive big results, you need to be able to move super fast, ship new content, experiment, learn and iterate. But most content management systems just aren't built for this. Your content teams wrestle with rigid interfaces as they build new pages. You spend endless time copying and pasting across pages and recreating content for other channels and applications. And their ideas for new experiments are squashed when developers can't build them within the constraints of outdated tech. Forward-thinking companies like Figma, Amplitude, Loom, Riot Games, Linear, and more use Sanity to build content growth engines that scale drive innovation and accelerate customer acquisition. With Sanity, your team can dream bigger and move faster. As the most powerful headless CMS on the market, you can tailor editorial workflows to match your business, reuse content seamlessly across any page or channel, and bring your ideas to market without developer friction. Sanity makes life better for your whole team. It's fast for developers to build with, intuitive for content managers, and it integrates seamlessly with the rest of your tech stack. Get started with Sanity's generous free plan. And as a Lenny's Podcast listener, you can get a boosted plan with double the monthly usage. Head over to sanity.io/lenny to get started for free. That's sanity.io/lenny.

中文翻译:

本集节目由 Sanity 赞助。你的网站是增长引擎的核心。为了让这个引擎产生巨大的成果，你需要能够极速行动、发布新内容、进行实验、学习并迭代。但大多数内容管理系统（CMS）并非为此设计的。你的内容团队在构建新页面时不得不与僵化的界面作斗争。你花费大量时间在页面间复制粘贴，并为其他渠道和应用重新创建内容。当开发人员无法在陈旧技术的限制下构建新实验时，他们的创意就会被扼杀。像 Figma、Amplitude、Loom、Riot Games、Linear 等具有前瞻性的公司都在使用 Sanity 来构建内容增长引擎，以实现规模化、驱动创新并加速客户获取。有了 Sanity，你的团队可以拥有更大的梦想并行动得更快。作为市场上最强大的无头 CMS（Headless CMS），你可以定制编辑工作流以匹配你的业务，在任何页面或渠道无缝复用内容，并在没有开发阻力的情况下将创意推向市场。Sanity 让整个团队的生活变得更美好。它对开发人员来说构建速度快，对内容经理来说直观易用，并且能与你的其他技术栈无缝集成。从 Sanity 慷慨的免费计划开始吧。作为 Lenny's Podcast 的听众，你可以获得每月使用量翻倍的增强版计划。前往 sanity.io/lenny 免费开始。地址是 sanity.io/lenny。

[00:03:26] [Lenny]

English:

This episode is brought to you by Hex. If you're a data person, you probably have to jump between different tools to run queries, build visualizations, write Python, and send around a lot of screenshots and CSV files. Hex brings everything together. Its powerful Notebook UI lets you analyze data in SQL, Python, or no code in any combination, and work together with live multiplayer and version control. And now Hex's, AI tools can generate queries and code, create visualizations, and even kickstart a whole analysis for you all from natural language prompts. It's like having an analytics co-pilot built right into where you're already doing your work. Then when you're ready to share, you can use Hex's drag and drop app builder to configure beautiful reports or dashboards that anyone can use. Join the hundreds of data teams like Notion, AllTrails, Loom, Mixpanel, and Algolia using Hex every day to make their work more impactful. Sign up today at hex.tech/lenny to get a 60 day free trial of the Hex team plan. That's hex.tech/lenny.

中文翻译:

本集节目由 Hex 赞助。如果你是一名数据人员，你可能需要在不同的工具之间跳转来运行查询、构建可视化、编写 Python，并发送大量的截图和 CSV 文件。Hex 将一切整合在一起。它强大的 Notebook 界面允许你以 SQL、Python 或无代码的任何组合方式分析数据，并通过实时多人协作和版本控制共同工作。现在，Hex 的 AI 工具可以根据自然语言提示生成查询和代码、创建可视化，甚至为你启动整个分析过程。这就像在你工作的地方内置了一个分析副驾驶（Co-pilot）。当你准备好分享时，可以使用 Hex 的拖拽式应用构建器来配置任何人都能使用的精美报告或仪表板。加入 Notion、AllTrails、Loom、Mixpanel 和 Algolia 等数百个数据团队的行列，每天使用 Hex 让工作更具影响力。立即在 hex.tech/lenny 注册，获得 Hex 团队计划的 60 天免费试用。地址是 hex.tech/lenny。

[00:04:31] [Lenny]

English:

Ramesh, thank you so much for being here. Welcome to the podcast.

中文翻译:

Ramesh，非常感谢你能来。欢迎来到本播客。

[00:04:35] [Ramesh Johari]

English:

Thanks so much for having me, Lenny. It's great to be on.

中文翻译:

非常感谢邀请我，Lenny。很高兴能参加。

[00:04:38] [Lenny]

English:

It's great to have you on. A big thank you to Riley Newman for connecting us. Riley was the first data scientist at Airbnb and head of data science at Airbnb, and that role is actually a really good microcosm of what we're going to be focusing on in our chat today. We're going to get super nerdy about marketplaces, and experimentation, and data. I know that's your jam. Are you ready to dive in?

中文翻译:

很高兴你能来。非常感谢 Riley Newman 为我们牵线搭桥。Riley 是 Airbnb 的第一位数据科学家，也是 Airbnb 的数据科学负责人，那个角色实际上是我们今天聊天重点的一个缩影。我们将深入探讨关于市场、实验和数据的硬核内容。我知道这是你的拿手好戏。准备好开始了吗？

[00:05:00] [Ramesh Johari]

English:

I really am. Yeah. And I actually want to thank Riley too. I got to know Riley when I was at oDesk first as a research scientist, and then I directed their data science team. This is way back in 2012, and I was looking around for people who are experts on how we think about data and marketplaces, and Riley Newman came up and so I invited him to come talk to us at oDesk, and we've stayed in touch since then. Those were early days of where this industry was, and I've had a kind of lengthy career now thinking about those kinds of problems. So I'm pretty excited to talk about it with you.

中文翻译:

我准备好了。我也想感谢 Riley。我在 oDesk 工作时认识了他，当时我先是担任研究科学家，后来领导他们的数据科学团队。那是远在 2012 年的事了，当时我正在寻找研究数据和市场思考方式的专家，Riley Newman 出现了，于是我邀请他来 oDesk 给我们做分享，从那以后我们就一直保持联系。那是这个行业发展的早期阶段，到现在为止，我在思考这类问题上已经有了相当长的一段职业生涯。所以我非常兴奋能和你探讨这些。

[00:05:37] [Lenny]

English:

Let's start broad and set a little foundation. You have a really interesting way to describe what a marketplace business even is. So Ramesh, what is a marketplace business, and also why is data so important and such an integral part of building a successful marketplace business?

中文翻译:

让我们从宏观开始，打好基础。你对“什么是双边市场业务”有一种非常有趣的描述方式。那么 Ramesh，什么是双边市场业务？此外，为什么数据如此重要，并且是构建成功市场业务不可或缺的一部分？

[00:05:53] [Ramesh Johari]

English:

It's interesting when people sit down and think about, say Airbnb, what does Airbnb sell? Average person is like, "That's pretty obvious. Airbnb sells rooms. I go there to book a room I want to stay at," right? Other people say, "What does Uber sell? Uber sells me rides. I'd use Uber when I need to get a ride from somewhere to somewhere else." And in some sense, you're not wrong. I mean, you go there. That's a platform to get these things. But that's not what the platform is selling. That's a really important distinction. There are people on the platform that are selling that to you. The hosts on Airbnb are selling you listings. The drivers on Uber are selling you rides. But Uber and Airbnb are selling you the taking away of something, which is a weird thing to think about. What they're taking away is the friction of finding a place to stay. They're taking away the friction of finding a driver.

In economics, we call those things transaction costs. When you take econ 1, you learn about markets and how supply meets demand, and we get prices out of that. But what you don't learn until econ 201 is that

markets don't always work. And one of the reasons markets don't always work is because we have what are called market failures due to the presence of these kinds of friction. So what's a market failure? It's that Lenny wants to get from Palo Alto to Burlingame and he can't do it. Why can't he do it? He doesn't have anyone to drive him. Well, why doesn't he just call someone to drive him? Well, who's he supposed to call? Who are those people? Are they out there? Are they willing to drive him right now, right at 10:00 AM on a Friday? Are they willing to take him somewhere? When I want to stay somewhere when I'm traveling, a friction is, who's willing to give me their room? I mean on principle, there's people who are willing to let me stay in their living room, but I don't know who they are.

So those are frictions, and what the marketplaces are selling you is taking the friction away. That's what you're paying them for. And it's an important observation, because what that means is the marketplace's customers aren't just the people buying the rides, they're buying the listings. Actually, the hosts are Airbnb's customers, and the drivers are also Uber's customers. So both sides of the marketplace are the customers of the platform. Both sides depend on the platform to help the platform take that friction away. Because just like you want a place to stay or you want a ride, the driver is at Uber because he wants to earn money by taking people places. And the host is on Airbnb because they want to earn money by selling their listing.

I think this concept that we're making money by taking transaction costs away is such a fundamental idea that's misunderstood around marketplaces. That when you're an entrepreneur starting a marketplace or thinking about your business model, I think you can be wildly off if you forget that that's the thing that's fundamentally your value proposition. And then you asked about the role of data, and more broadly data science in marketplaces.

So it's an interesting thing, right? The example I always love to give are the ancient Agoras in Greece or Trajan's Market in Rome. When you look at pictures of these things, what really stands out to me is the rock. I mean, these things are made of stone. It's not like you were going to move a booth from one place to another place without moving a lot of rock from one place to another place. So you flash forward to 2023, and here we are with technology undergirding pretty much every kind of commerce now. And it means we can architect and re-architect the marketplace kind of on the fly, and we really are doing it all the time. And these frictions that are getting taken away, they're getting taken away because of data and data science. So I really want to highlight three pieces of this for people, which I want you to think of them as a cycle. But to start with, let's just lay them out one at a time.

One of them is finding people to match with. So that's the problem of, "I want to stay somewhere. Who is out there, who's willing to let me stay with them on a given timeframe?" And then if I'm a host, I have a listing. Who is out there, who's willing to stay at my place when I have it available? So that's finding matches. Then there's making the match. And so here, going back to my time at oDesk, a big problem that we dealt with there was if I've got multiple applicants to my job, who should I hire? Who should I interview? It's a common problem we face in the real world, but now it's all remote. I don't meet these people in person. All I've got is this application they submitted to me. I need help triaging that. Okay? So that's helping make a match out of possible partners you can match with.

And then finally, we make matches. Well, what do the matches tell us, right? I mean, if you stay somewhere in Airbnb, you learn something about the host, you learn something about the listing. The host learns about you too. And that's all information that the marketplace should feed back in. So this is where we get to rating systems and feedback systems, even passive data collection, right? Did you leave your booking before you were supposed to leave? Well, maybe that's a sign that something didn't quite work out the way you wanted to work out. So that's passive data collection. Did you leave five stars? That's active data collection. Get all this back in, and what does that do? Well, that lets us do a better job finding potential matches and make potential matches in the future. Every single thing I just said, finding potential matches, making matches, and then learning about those matches, and then cycling back

again, that is the data science in marketplaces. And I feel like every marketplace that you could think of in any vertical has those three problems to deal with and relies on algorithms in data science to help them solve it. And in turn, that is I think really the underpinning of taking those frictions away.

中文翻译:

很有意思，当人们坐下来思考，比如 Airbnb 时，Airbnb 卖的是什么？普通人会说：“这很明显，Airbnb 卖房间。我去那里预订我想住的房间，”对吧？另一些人会说：“Uber 卖什么？Uber 卖给我行程。当我需要从一个地方去另一个地方时，我会用 Uber。”从某种意义上说，你没说错。我的意思是，你去那里，那是一个获取这些东西的平台。但那并不是平台在“卖”的东西。这是一个非常重要的区别。平台上有人在向你出售这些东西。Airbnb 的房东在向你出售房源。Uber 的司机在向你出售行程。但 Uber 和 Airbnb 卖给你的是“消除某种东西”，这听起来很奇怪。他们消除的是寻找住宿的“摩擦力”(Friction)。他们消除的是寻找司机的摩擦力。

在经济学中，我们称之为“交易成本”(Transaction costs)。当你上经济学 101 时，你学到的是市场以及供给如何满足需求，并由此产生价格。但直到经济学 201 你才会学到，市场并不总是有效的。市场失灵的原因之一就是存在这些摩擦力。什么是市场失灵？比如 Lenny 想从帕罗奥图去伯灵格姆，但他去不了。为什么？因为没人载他。那他为什么不打电话叫人载他？他该打给谁？那些人在哪？他们现在愿意载他吗，就在周五上午 10 点？他们愿意带他去那个地方吗？当我想在旅行时找个地方住，摩擦力在于：谁愿意把房间给我？原则上，有人愿意让我住在他们的客厅里，但我不知道他们是谁。

所以这些就是摩擦力，而双边市场卖给你的是“消除摩擦力”。这就是你付钱给他们的原因。这是一个重要的观察，因为这意味着市场的客户不仅仅是买行程或订房的人。实际上，房东也是 Airbnb 的客户，司机也是 Uber 的客户。市场的两端都是平台的客户。两端都依赖平台来帮助消除摩擦力。因为就像你想找地方住或想搭车一样，司机在 Uber 是因为他想通过载人赚钱，房东在 Airbnb 是因为他们想通过出租房源赚钱。

我认为这种“通过消除交易成本来赚钱”的概念是一个非常基础的想法，但在双边市场领域常被误解。当你作为一个创业者开始建立一个市场或思考你的商业模式时，如果你忘记了这才是你核心的价值主张，你可能会错得离谱。接着你问到了数据，以及更广泛的数据科学在市场中的作用。

这很有趣，对吧？我最喜欢举的例子是古希腊的集市 (Agoras) 或罗马的图拉真市场 (Trajan's Market)。当你看到这些地方的照片时，最让我印象深刻的是石头。我是说，这些东西是石头做的。你不可能在不搬动大量石头的情况下，把一个摊位从一个地方挪到另一个地方。快进到 2023 年，现在的技术支撑着几乎每一种商业形式。这意味着我们可以随时随地架构和重构市场，而我们确实一直在这么做。这些被消除的摩擦力，正是通过数据和数据科学被消除的。我想为大家强调其中的三个环节，你可以把它们看作一个循环。

第一是“寻找匹配对象”。这就是：“我想住在某个地方，谁在那里，谁愿意在特定时间内让我住？”如果我是房东，我有一个房源：“谁在那里，谁愿意在我有空房的时候入住？”这就是寻找匹配。

第二是“达成匹配”。回到我在 oDesk 的时光，我们处理的一个大问题是：如果我的职位有多个申请人，我该雇佣谁？我该面试谁？这是现实生活中常见的问题，但现在全是远程的。我没见过这些人。我手里只有他们提交的申请。我需要帮助来筛选 (Triage)。所以，这是在可能的合作伙伴中帮助达成匹配。

最后是“从匹配中学习”。我们达成了匹配，那么匹配告诉了我们什么？如果你住在 Airbnb 的某个地方，你会了解房东，了解房源。房东也会了解你。这些都是市场应该反馈回来的信息。这就是评价系统和反馈系统发挥作用的地方，甚至包括被动数据收集。比如：你是不是在预定结束前就离开了？这可能是一个信号，表明某些事情没按预期发展。这是被动数据收集。你留了五星好评吗？这是主动数据收集。

把所有这些信息反馈回来，有什么用？这能让我们在未来更好地寻找潜在匹配并达成匹配。我刚才说的每一件事——寻找潜在匹配、达成匹配、学习匹配结果，然后再次循环——这就是双边市场中的数据科学。我觉得你能想到的任何垂直领域的市场都要处理这三个问题，并依赖数据科学中的算法来帮助解决。反过来，我认为这正是消除摩擦力的基石。

English:

Many founders try to start a marketplace business, think about marketplace opportunities where they don't exist. And there's often these recurring failures of types of marketplaces that just don't work in an area. I was just writing a couple ideas down while you were chatting like cleaners, getting cleaners as a marketplace doesn't seem to work ever. Car wash, there's a classic failure too. Getting tasks done for you on demand as a marketplace seems to not often work. So this might be too big of a question, but I'm just curious if anything comes up of when someone is starting a marketplace or thinking about starting a marketplace business, what do you find are the most common flaws in, this is probably not going to work as a marketplace?

中文翻译:

许多创始人尝试启动双边市场业务，在不存在机会的地方寻找机会。在某些领域，某些类型的市场总是反复失败。刚才你聊天时我记下了几个点，比如清洁工，把清洁工做成双边市场似乎从来没成功过。洗车也是一个经典的失败案例。按需完成任务的市场似乎也经常行不通。这可能是一个太大的问题，但我很好奇，当有人开始或考虑开始一个市场业务时，你发现最常见的缺陷是什么，即哪些迹象表明“这可能不适合做成双边市场”？

[00:12:01] [Ramesh Johari]

English:

That is such a fantastic question, and I want to preface what I say with a couple of comments. So one of them is that I've worked with a lot of different marketplace companies, but anything I say is pertaining to something more sensitive. I may not name the company over the course of the podcast. But the other more important thing I want to say is that I'm a professor at Stanford, and there's a reason I'm not a successful scaled entrepreneur of marketplaces, and that's because I probably haven't unlocked the key to exactly the question you asked. Nevertheless, I have some thoughts on it.

The most important one is this. What I've found talking to people who want to start what they think is a marketplace is that they think too much about a marketplace before they're a marketplace. That in my view is the biggest failure mode. You mentioned specific things, cleaners. I wonder about that, right? Is it something about the cleaning industry? It possibly is. I don't claim to be an expert on the microeconomics of the cleaning industry. But often it's not that, it's that I thought I was building a marketplace from the beginning, and that's not the way the world works. So I'll give you one vignette of this that I really like, and that's UrbanSitter.

So first, UrbanSitter is a babysitting marketplace. We can talk about their whole life story, but I think what's most interesting is really the early days. And in the early days, what I found interesting, the way I found out about them actually is that we were stuck looking for some help. And I found out about this new platform where the clever thing was when you used to hire a babysitter, it's like pre Venmo days, you needed cash on hand. Because when the babysitter's done at the end of the day, they're usually high school suits or something. They want to get paid. They're not going to take your IOU, that you'll send them some check in the mail the next day. And unfortunately, you often don't have cash. They don't take credit card. They're high school students. That was an incredible friction to address, which is literally just we accept credit card payments for babysitting. That's it, right?

Now from there, what happened is they took advantage of Facebook networks between parents and babysitters to build trusted introductions. So let's say my sitter wasn't available. I get to know sitters in the Facebook network of that sitter. And once they overcame that first thing to get some liquidity onto their platform, they could move towards asking, how do I solve for these frictions that I talked about earlier? How do I solve for helping people find potential matches? How do I solve for people making those matches, right? You can't do that when you don't have liquidity on your platform. It's silly to tell

someone, "Hey, I'm really going to help you find all those drivers out there, even though I only have three drivers on my platform." That's not a friction you're solving for.

In their example, as they evolved, they actually shifted their monetization model away from billing specifically for this friction of allowing you to pay with credit cards, instead to now billing for how you were interviewing and contacting sitters. They had a two-part plan for that. One with a pay as you go menu, one with a more of a subscription option. But the key thing was either way, what you were paying for now was finding potential babysitters, not paying them with a credit card. That wasn't the key thing anymore.

So what's the moral there? The moral is a marketplace business never starts as a marketplace business, because what we think of as a marketplace business is something which at scale is removing the friction of the two sides finding each other. But when you start, you don't have that scale. So when you start, you had better be thinking, "What's my value proposition in a world in which I don't have that scaled liquidity on both sides?" And that's bespoke. It means different things. And in the case of oDesk, where I started, that initial thing was that remote work is a weird thing, because basically you've somehow got to know that this person who you're not next to is doing what you're asking them to do. And so the initial value proposition of oDesk was to provide tools for workers to verify they were working the hours and doing the things that they said they were doing, screenshots and various kinds of tracking.

And then in return for that, to be able to provide guarantees on both sides. So now the workers could say, "Hey, I worked what I said so I should get paid." And the employers could say, "I actually see that you worked what I said. And so I feel comfortable that I got what I paid for." That was the initial value proposition, is resolving a trust issue at a remote scale. At that point, liquidity isn't the game. It's asking, what's a problem that people in this space are facing that I can deal with when I'm not a scaled marketplace? So again, with the cleaning industry, I can comment on that from personal experience, but otherwise, I think that's the way I would think about it. It's almost never about building a marketplace when you're building a marketplace.

中文翻译:

这是一个非常棒的问题。在回答之前，我想先做两点说明。第一，我曾与许多不同的市场公司合作过，但我说的任何内容如果涉及敏感信息，我可能不会在播客中点名具体公司。第二点更重要，我是斯坦福大学的教授，我之所以不是一个成功的规模化市场创业者，是有原因的，那就是我可能还没有解开你刚才问的那个问题的关键钥匙。尽管如此，我还是有一些想法。

最重要的一点是：我发现那些想创办“他们认为的市场”的人，在真正成为市场之前，考虑了太多关于“市场”的事情。在我看来，这是最大的失败模式。你提到了具体的事情，比如清洁工。我也在想，是因为清洁行业的某些特性吗？很有可能。我不自称是清洁行业微观经济学的专家。但通常原因不在于此，而在于“我从一开始就认为我在建立一个市场”，而世界并不是这样运作的。我给你举一个我很喜欢的例子，就是 UrbanSitter。

首先，UrbanSitter 是一个保姆（Babysitting）市场。我们可以聊他们的整个故事，但我认为最有趣的是早期。在早期，我发现他们很有趣的一点是——我之所以知道他们，是因为当时我们急需找人帮忙。我发现了这个新平台，它的聪明之处在于：在 Venmo 出现之前的日子里，你雇保姆需要准备现金。因为保姆干完活通常是在傍晚，他们大多是高中生之类的，他们想拿钱走人。他们不会接受你的欠条，或者让你第二天邮寄支票。但不幸的是，你经常没有现金。他们又不收信用卡，因为他们只是高中生。这是一个非常大的摩擦力，而 UrbanSitter 解决的摩擦力仅仅是：我们接受信用卡支付保姆费。就这么简单，对吧？

从那里开始，他们利用父母和保姆之间的 Facebook 社交网络来建立信任推荐。比如，如果我的保姆没空，我可以认识那个保姆 Facebook 网络里的其他保姆。一旦他们克服了第一个障碍，让平台上有了流动性（Liquidity），他们就可以转向解决我之前提到的那些摩擦力：如何帮助人们找到潜在的匹配？如何帮助人们达成匹配？当你平台上没有流动性时，你是做不到这些的。如果你告诉别人：“嘿，我会帮你找到外面所有的司机，”尽管你平台上只有三个司机，那是很愚蠢的。那不是你在解决的摩擦力。

在他们的例子中，随着发展，他们实际上将盈利模式从“为信用卡支付这一摩擦力收费”转变为“为你面试和联系保姆收费”。他们有一个两部分的计划：一个是按次付费，一个是订阅制。但关键在于，无论哪种方式，你现在支付的是“寻找潜在保姆”的费用，而不是“用信用卡付钱”。后者已经不再是核心了。

那么这里的寓意是什么？寓意是：一个市场业务从来不是以“市场业务”开始的。因为我们所认为的市场业务，是在规模化之后消除两端寻找彼此的摩擦力。但当你开始时，你没有那个规模。所以当你开始时，你最好思考：“在一个两端都没有规模化流动性的世界里，我的价值主张是什么？”这是定制化的，意味着不同的东西。以我开始工作的 oDesk 为例，最初的东西是：远程工作是一件很奇怪的事，因为你必须以某种方式知道那个不在你身边的人确实在做你要求他做的事。所以 oDesk 最初的价值主张是为工人提供工具，证明他们确实工作了那么长时间并做了那些事，比如截图和各种追踪。

作为回报，平台能为两端提供保证。工人可以说：“嘿，我按要求工作了，所以我应该拿到钱。”雇主可以说：“我确实看到你按要求工作了，所以我付钱付得放心。”这就是最初的价值主张：在远程规模下解决信任问题。在那一点上，流动性并不是重点。重点是问：这个领域的人们面临的、我在还不是规模化市场时就能处理的问题是什么？所以回到清洁行业，我可以从个人经验出发评论，但除此之外，我认为这就是思考方式。当你建立一个市场时，重点几乎从来不在“建立一个市场”。

[00:16:43] [Lenny]

English:

That's very similar to the advice I always give marketplace founders, is 90% of your problems are going to be non marketplace specific problems. They're going to be the same problems any startup is going to have. How do I grow? It's going to be the same things you need to do.

中文翻译:

这和我经常给市场创始人的建议非常相似：你 90% 的问题都将是非市场特有的问题。它们是任何初创公司都会遇到的相同问题。比如“我该如何增长？”你需要做的事情是一样的。

[00:16:58] [Ramesh Johari]

English:

So one thing you said was, "That's what you tell marketplace founders." I mean, something I've actually pressed hard on in my own way of thinking about this is that maybe we shouldn't talk about the concept of a marketplace founder. Really there's founders. And I think every entrepreneur... I mean one way to think about it, right? It's very hard to think about a human business endeavor that has not been disrupted by the potential for transactions to take place online. And if that's the case, it means literally any founder is a marketplace founder. It'll be a choice they make after they grow as to whether they want to build a platform. I mean, to take a very hot recent example, no one in their right mind would've thought of OpenAI as a marketplace, but OpenAI is a marketplace now. They may not want to call themselves a marketplace, but they have plugins. The plugins are flooding that platform. People have played with it. It's not an easy thing to find the plugin you need for what you want to do. And that really is a two-sided thing now. There's the plugin creators and there's the users. And they may believe it, they may not believe it, but they are a marketplace.

So I think a different way to think about it is every founder is a marketplace founder. It's going to be a choice they want to make for themselves of whether they want to become that platform. That's I think one. And two is because that's the case, I think one of the other challenges I find founders struggle with is you don't want to overcommit your future. And what I mean by that is that you're building up trust, and you're building up a sense of what kind of business you are in your early days. If you believe that this kind

of platform future awaits you, or market platform future awaits you, there may be choices you're making early on that are tying your hands later.

A great example of this is when oDesk started, it was because the tools they were providing were for ongoing monitoring of work. It's a very natural thing to say, "We will just take a constant fraction of the dollars that cross the platform." That all works well and good until after you become mature. Some of these relationships between worker and employer last a long time, and most of the value was generated now not so much because they're able to track each other, because the trust is now there, but because they found each other, because they're able to build that relationship through oDesk. That meant that the longer that goes on, the less value the platform is adding into that relationship, but you're still pulling 10% of all the dollars. So what does that lead to? A word that most marketplace CEOs know well is disintermediation, which is where you were intermediating between the two parties, and now disintermediation means that essentially they're like, "Hey, we don't need you anymore."

My favorite example is we had some stuff delivered from IKEA by a Thumbtack worker once, and my wife is like, "Oh, thanks a lot. You're so reliable." He's like, "Hey, great. Here's my business card. Ever need me again? Just call the number on the back." And that was it. Thumbtack got their one lead gen, and then we didn't need the platform anymore. And I think this issue for oDesk meant that after they merged with Elance and became Upwork, they had to think a little bit about, "Okay, what's the monetization strategy we want to use? How do we address this issue that longer term relationships may disintermediate? And does that mean you need a pricing plan that actually takes that into account?" So early commitments in this case to a particular pricing scheme, particular monetization, can really tie your hands as you then realize later you actually are a platform.

中文翻译:

你刚才说：“这就是你告诉市场创始人的话。”我想在我自己的思考方式中强调一点：也许我们不应该讨论“市场型创始人”这个概念。实际上只有“创始人”。我认为每一个企业家……我的意思是，可以这样想：很难想象有什么人类商业活动还没有被在线交易的可能性所颠覆。如果是这样的话，这意味着字面意义上的任何创始人都是市场型创始人。这只是他们在成长之后做出的一个选择，即他们是否想要构建一个平台。举一个最近很火的例子，正常人以前绝不会把OpenAI看作一个市场，但OpenAI现在就是一个市场。他们可能不想称自己为市场，但他们有插件（Plugins）。插件正充斥着那个平台。人们已经体验过了，要找到你需要的插件并不容易。这现在确实是一个双边的事物：有插件创建者，也有用户。不管他们信不信，他们就是一个市场。

所以，我认为另一种思考方式是：每个创始人都是市场型创始人。是否要成为那个平台，是他们自己要做的选择。这是第一点。第二点是，正因为如此，我发现创始人面临的另一个挑战是：你不想对你的未来过度承诺。我的意思是，你在早期建立信任，建立一种“你是什么样的公司”的感觉。如果你相信这种平台化的未来在等着你，那么你早期做出的一些选择可能会在以后束缚你的手脚。

一个很好的例子是oDesk刚开始时，因为它提供的工具是用于持续监控工作的。很自然地会说：“我们将从流经平台的每一分钱中抽取固定比例。”这在成熟之前运作良好。但当工人与雇主之间的某些关系持续很长时间后，大部分价值的产生不再是因为他们能互相追踪（因为信任已经建立了），而是因为他们找到了彼此，因为他们能通过oDesk建立这种关系。这意味着关系持续得越久，平台在其中的增值就越少，但你仍然抽取10%的费用。这会导致什么？大多数市场CEO都很熟悉的一个词是“去中介化”（Disintermediation），即你原本在两方之间做中介，现在去中介化意味着他们基本上会说：“嘿，我们不再需要你了。”

我最喜欢的例子是，有一次我们请Thumbtack的工人送IKEA的东西，我妻子说：“噢，太感谢了，你真可靠。”他说：“嘿，太棒了，这是我的名片。以后还需要我吗？直接打背面的电话就行。”就这样。Thumbtack只获得了一次获客（Lead gen），然后我们就再也不要这个平台了。我认为对于oDesk来说，这意味着在他们与Elance合并成为Upwork之后，他们必须思考：“好吧，我们想用什么样的盈利策略？我们如何解决长期关系可能去中介化的问题？这是否意味着你需要一个真正考虑到这一点的定价方案？”在这种情况下，早期对特定定价方案或特定盈利模式的承诺，当你后来意识到自己其实是一个平台时，真的会束缚你的手脚。

[00:20:26] [Lenny]

English:

I really like this message. It makes me think about Substack actually, which started as just a platform for newsletter writers. And then they're like, "How do we make this more valuable?" Because they take a cut of everyone's revenue. And they've actually invested heavily on helping drive demand to writers, for example, me. And at this point, over 80% of my subscribers come from Substack's network. And so they've built this marketplace element exactly as you're describing, where they just found, "Here's a pain point, writers need more subscribers. How do we help them drive subscribers?" So they figured out all these ways to create demand.

中文翻译:

我很喜欢这个观点。这让我想到了 Substack，它最初只是一个为时事通讯（Newsletter）作者提供的平台。然后他们想：“我们如何让它更有价值？”因为他们从每个人的收入中抽成。他们实际上投入了大量精力帮助作者引入需求，比如我。目前，我超过 80% 的订阅者来自 Substack 的网络。所以他们正是按照你描述的方式构建了这个市场元素，他们发现了：“这是一个痛点，作者需要更多订阅者。我们如何帮助他们获取订阅者？”于是他们想出了各种创造需求的方法。

[00:20:58] [Ramesh Johari]

English:

That's a really positive story, where they managed to actually expand the frontier of their business by enabling that network. For every one of those, there's unfortunately a lot of negative stories. I mean, one that I think is very painful is how eBay had a lot of challenges with its seller community as it introduced more and more fine-grained sources of fees. And I think a lot of that, I mean there's many, many treatises at this point written on eBay, and their history, and how they got to the point that they're at. But I think one kind of simple thing I do want people to think about there is that the sellers on eBay who had matured with the platform, who had grown with it, had come to develop certain expectations about what their lives on that platform would look like. And it's understandable, because a lot of these businesses, they had built their livelihood on that platform. That was their entire business. So when you now reach in and you say, "I'm going to completely change the rules of the game in which your business model operates," from the perspective of those sellers, that's a breaking of a social contract that's been developed over a very long time. So I love the Substack example, because that's like, "Hey, let me amplify our social contract." But I think for every one of those, there's an eBay warning sign that you can trap yourself a little bit.

中文翻译:

这是一个非常正面的故事，他们通过激活那个网络成功扩展了业务边界。但不幸的是，每一个这样的故事背后都有很多负面的案例。我认为一个非常痛苦的例子是 eBay，随着它引入越来越多细化的费用来源，它与卖家社区之间产生了巨大的挑战。关于 eBay 的历史以及他们如何走到今天这一步，已经有很多论著了。但我希望人们思考的一点是：eBay 上的卖家是伴随着平台成熟和成长的，他们对在这个平台上的生活已经形成了一定的预期。这是可以理解的，因为很多这类商家是靠这个平台谋生的，那是他们的全部业务。所以当你现在介入并说：“我要彻底改变你商业模式运作的游戏规则”时，从卖家的角度来看，这是在破坏长期以来形成的“社会契约”。所以我喜欢 Substack 的例子，因为那是“嘿，让我加强我们的社会契约”。但我觉得每一个那样的例子，都有一个 eBay 式的警示信号，提醒你可能会陷入困境。

[00:22:24] [Lenny]

English:

Just to close a loop on this really, I think important point, a lot of people listening to this are probably, "I'm a marketplace founder. I'm building a marketplace," are going to hear this and be like, "Oh shit, maybe I need to rethink how I think about what I'm doing." What would be your piece of advice to people like that? Is it focus on the friction point and it may be a marketplace solution, it may be a managed marketplace, it may be you own the supply? Is that the advice, or what would your advice be to someone that's like, "I'm building a marketplace"? How should they reframe their thinking?

中文翻译:

为了给这个非常重要的观点收个尾，很多听众可能会想：“我是一个市场创始人，我正在建立一个市场。”听到这些后可能会觉得：“噢，糟糕，也许我需要重新审视我正在做的事情。”你会给这些人什么建议？是专注于摩擦点，而解决方案可能是一个市场，也可能是一个托管型市场（Managed marketplace），甚至可能是你拥有供应端？这是你的建议吗？或者你会对那些说“我正在建立一个市场”的人说什么？他们应该如何重构思维？

[00:22:54] [Ramesh Johari]

English:

Let's go back to kind of thinking about this concept of a marketplace of reducing friction. So the litmus test I like to give to someone who claims to me that they're building a marketplace business or they're a marketplace founder is do you have what I would call scaled liquidity on both sides of your platform? What does scaled liquidity mean? What it means in lay terms... And by the way, I am a data scientist, and I love to think about these quantitatively. But fundamentally, if it doesn't pass the smell test, then you don't have to keep going with the data science. The smell test is scaled liquidity asks, "Do I have a lot of buyers and a lot of sellers on my platform, or do I only have one of these two, or do I have neither?" If you don't have both, you could call yourself whatever you want to call yourself, but at this moment in time, you're not a marketplace. If you have one, congratulations. You've won the game on one side of the market. And now you if you want, you have a choice point. You can lean into growth on the side that you're doing well with. You got a ton of users, ton of buyers? Great. Lean into it, get more buyers. That's one option. There's no shame in not being a marketplace. Scaling a business is scaling a business. If that's the way to do it, do it.

If you decide you want to be a marketplace, then at that moment when you've got a lot of buyers, but not a lot of sellers, or a lot of sellers, but not a lot of buyers, the choice you're facing is, how do I take advantage of having that one side scaled to attract the other side? We can talk more about that, but there's a lot of ways to kind of hack that, to think about how... So to take Uber as an example, they would walk into a new city, and one thing that Uber was commonly known for doing this was back in the days when really Uber Black was the only service is they just hand out coupons for free rides at events, parties, things like that, to take people home. And that was a way of saying, "Hey, we're subsidizing the drivers in the city. That's our scaled side. Now we're going to use that subsidized driver base to attract riders." So that's like, how do you get that flywheel going? And again, many people have written about how to take liquidity, scaled liquidity on one side, and use it to attract the other side.

If you don't have either side, don't worry about it. Don't worry about being a marketplace. Worry about scaling one side. And in that world, it opens your visibility up completely into the advice of many, many startup advisors. People who have advice not so much about scaling a marketplace, but about scaling a startup. And I want to say you got to let the ego go at that point. It's fine to articulate to people that your vision of the future is to be a platform or marketplace. As I said, virtually every business is going to have that option at some point in the modern tech enabled economy anyway. So you're not saying something people don't already know when you tell an advisor or an investor that. But I do think you need to be

humble enough at the starting point to recognize that there's no sense in talking about a marketplace if you don't have scaling on either side yet.

中文翻译:

让我们回到“市场是减少摩擦”这个概念。对于那些声称自己在建立市场业务或自己是市场创始人的人，我喜欢给他们做一个“试金石测试”：你的平台两端是否拥有我所说的“规模化流动性”（Scaled liquidity）？规模化流动性是什么意思？用通俗的话来说……顺便说一下，我是一名数据科学家，我喜欢定量地思考这些问题。但从根本上说，如果它没通过“直觉测试”（Smell test），那你就不必继续用数据科学分析了。直觉测试就是问：“我的平台上有很多买家和很多卖家吗？还是只有其中之一？或者两者都没有？”如果你两端都没有，你可以随你怎么称呼自己，但在此时此刻，你还是一个市场。如果你拥有一端，恭喜你，你在市场的一端赢得了比赛。现在如果你愿意，你面临一个选择点。你可以继续深耕你做得好的那一端。你有大量用户、大量买家？太棒了，继续投入，获取更多买家。这是一种选择。不做一个市场并不可耻。规模化业务就是规模化业务，如果那是正确的方法，那就去做。

如果你决定要做一个市场，那么当你有很多买家但卖家不多，或者有很多卖家但买家不多时，你面临的选择是：我如何利用已经规模化的一端来吸引另一端？我们可以深入探讨，有很多“黑客”方法可以思考如何做到这一点。以 Uber 为例，当他们进入一个新城市时（在只有 Uber Black 服务的年代），他们常做的一件事就是在活动、派对等场合发放免费乘车券。这是一种表达方式：“嘿，我们正在补贴这个城市的司机，这是我们规模化的一端。现在我们要利用这个受补贴的司机群体来吸引乘客。”这就是如何让飞轮转动起来。关于如何利用一端的规模化流动性来吸引另一端，已经有很多人写过相关文章了。

如果你两端都没有，那就别担心。别担心自己是不是一个市场。担心如何规模化其中一端。在那个阶段，你可以完全采纳许多初创公司顾问的建议——那些关于如何规模化初创公司，而不是规模化市场的建议。我想说，在那一点上你必须放下自尊。向人们阐述你未来的愿景是成为一个平台或市场是没问题的。正如我所说，在现代技术驱动的经济中，几乎每家企业最终都会有这个选择。所以当你告诉顾问或投资者时，你并不是在说一些他们不知道的事情。但我认为你在起点时需要足够谦逊，意识到如果两端都还没规模化，谈论“市场”是没有意义的。

[00:25:52] [Lenny]

English:

And then it becomes a question of a business model, unit economics of, can I build say a DoorDash, not as a marketplace? Can I just hire a bunch of people delivering? Is this even possible in a different route?

中文翻译:

然后这就变成了一个商业模式和单位经济效益（Unit economics）的问题：我能不能建立一个像 DoorDash 这样的公司，但不作为市场？我能不能直接雇佣一群人来送货？通过不同的路径这是否可行？

[00:26:06] [Ramesh Johari]

English:

Yeah, that's a great point. One of the things I think that's useful for people to think about here that you're raising, at some level, it's kind of tied up I think with that question of whether I should have employees, or contract or freelance work on one side of the marketplace. And that's actually a pretty old question in economics. The way we talk about it often is a distinction between a market or a firm. And one of the interesting puzzles in economics, Ronald Coase is a famous economist who thought about this is, "Well, if markets are so efficient, why do we need firms? If markets are efficient at matching labor up with things that need to get done, why would you ever need a firm?" And that's one of the earliest recognitions that transaction costs are a real thing. And that's one of the things that firms are solving for.

And I love what you're saying because what it's recognizing is, "Hey, for your frictions, the best resolution to that might not be to have a marketplace. It might actually be to have very tightly controlled labor." A good example of this actually, Stitch Fix, I think one of the things that's cool about Stitch Fix is the experience that people had early on with stylists at Stitch Fix.

中文翻译:

是的，这是一个很好的观点。你提到的这一点对人们很有用，在某种程度上，我认为它与“我应该在市场的一端拥有员工，还是合同工或自由职业者”这个问题紧密相关。这实际上是经济学中一个相当古老的问题。我们通常将其表述为“市场”与“企业”的区别。经济学中有一个有趣的谜题，著名经济学家罗纳德·科斯（Ronald Coase）曾思考过：“如果市场如此高效，为什么我们需要企业？如果市场能高效地将劳动力与需要完成的事情匹配起来，为什么还需要企业？”这是对“交易成本是真实存在的”最早的认知之一。而这正是企业所要解决的问题之一。

我很喜欢你说的，因为它意识到：“嘿，对于你面临的摩擦力，最好的解决方案可能不是建立一个市场，而可能是拥有受严格控制的劳动力。”一个很好的例子是 Stitch Fix，我认为 Stitch Fix 很酷的一点是人们早期与 Stitch Fix 造型师的互动体验。

[00:27:13] [Lenny]

English:

I'm a happy customer, by the way. I think [inaudible 00:27:16].

中文翻译:

顺便说一下，我是他们的忠实客户。我觉得……（听不清）。

[00:27:15] [Ramesh Johari]

English:

Yeah, I think one of the great things about that experience is it felt magical to have someone who kind of got to know you. But that depends on a relationship that doesn't feel like a freelance relationship every single time you're going back. Another example that I would pull out is pretty much any healthcare platform. So for example, for physical therapy, it'd be weird if every time you went to a physical therapy platform, you just got randomly matched to whoever happened to be available then. So I think there's some curation that needs to happen of that relationship. Does that mean full employee? Maybe not. But it does mean you have to think a little bit about exactly as you brought up, what's the nature of curation of your labor pool?

中文翻译:

是的，我认为那种体验最棒的地方在于，有一个人能逐渐了解你，这感觉很神奇。但这取决于一种关系，这种关系不能让你每次回来都觉得像是在面对一个自由职业者。我举的另一个例子是几乎任何医疗平台。比如物理治疗，如果你每次去物理治疗平台，都只是被随机分配给当时恰好有空的人，那会很奇怪。所以我认为这种关系需要某种“策展”（Curation，即筛选与维护）。这是否意味着必须是全职员工？也许不是。但它确实意味着你必须思考，正如你提到的，你的劳动力池的策展性质是什么？

[00:28:01] [Lenny]

English:

Awesome. Okay, so let's come back to a point you made early on around the importance of data and the power of data in actually making your marketplace a lot more efficient and work more effectively. So say that you have a data scientist, or a data analyst, or someone that is helping you optimize your marketplace. Where do you often find the biggest leverage and opportunity for a data person to help you make your marketplace more effective?

中文翻译:

太棒了。好，让我们回到你早些时候提到的关于数据重要性的观点，以及数据在让市场更高效运作方面的力量。假设你有一名数据科学家、数据分析师或某人正在帮助你优化市场。你通常认为数据人员在哪些方面能发挥最大的杠杆作用和机会，从而让你的市场更有效？

[00:28:25] [Ramesh Johari]

English:

This is an incredible question, right? Because I think I could answer it a number of different ways. One question I think that's kind of basic, it's just what should this person be doing? And I'm going to actually evade that question a little bit. I'm going to give some examples of what they could do, but I feel like that's one where context matters a lot. So as an example, at ride-sharing or grocery delivery marketplaces, pricing means actually, what do you pay for that ride? Or what do you pay for that delivery? So that's actually the price that's set at the moment you actually place the order. Just to be clear, by the way, if you order from DoorDash, I don't mean the price of the restaurant. I mean, what do you pay to DoorDash, right? What's that fee? Is there a surcharge, because it's surge or whatever? Okay, so that's a thing, right?

But that's not really a thing in a marketplace where the platform's not setting the prices. So in Airbnb, really hosts are the ones who are in charge of setting prices for their listings. One answer to your question is, if I'm in a place like Uber, Lyft, DoorDash, I want to have good data scientists thinking about pricing. Because that seems like something which should be heavily dependent on the instantaneous state of supply and demand in my marketplace. So that's one type of answer is, well do I need data scientists working on pricing? Do I need data scientists working on search? Why search? Because maybe in my marketplace, finding the needle in the haystack is really the biggest, highest friction problem. So maybe I need a lot more data scientists saying about search. That's what I'm going to evade. Okay? I'm going to focus more on something completely different, which is just a more philosophical point about what a data scientist does.

中文翻译:

这是一个极好的问题。因为我觉得我可以从很多不同的角度来回答。一个基本的问题是：这个人应该做什么？我打算稍微回避一下这个问题。我会举一些他们可以做的例子，但我觉得这很大程度上取决于具体的语境。例如，在网约车或杂货配送市场中，“定价”实际上意味着你为那次行程或那次配送支付多少钱。这实际上是在你下单那一刻设定的价格。顺便澄清一下，如果你从 DoorDash 订餐，我指的不是餐厅的价格，而是你付给 DoorDash 的钱。那笔费用是多少？是否有加价（Surge）？好，这是一个方向。

但在平台不设定价格的市场中，这就不太适用了。比如在 Airbnb，房东才是负责设定房源价格的人。所以对你问题的一个回答是：如果我在 Uber、Lyft、DoorDash 这样的公司，我希望有优秀的数据科学家思考定价。因为这似乎应该高度依赖于市场中供应和需求的瞬时状态。所以这是一种答案：我需要数据科学家研究定价吗？我需要数据科学家研究搜索吗？为什么是搜索？因为也许在我的市场中，“大海捞针”才是最大、摩擦力最高的问题。所以也许我需要更多的数据科学家研究搜索。这就是我要回避的部分。好吗？我将更多地关注一些完全不同的东西，即关于数据科学家工作的更具哲学意义的一点。

English:

So in a lot of companies today, especially, a main thing that you ask data scientists to do is build what's called a machine learning model. Now, machine learning model even already can mean a lot of things to a lot of different people. I'm going to focus on something very concrete. You're asking them to predict something. When I started at oDesk, this is in 2012, one of the funny things about me is I started at oDesk because I'd had a academic career up to that point in about 10 years, just building mathematical models of things. I was not really very much of a data scientist up to that point. What I expected would happen is I'd go to industry and I'd be told, "Hey, look how important data is." And definitely my eyes were opened.

And one of the first things I was asked to think about is, well, okay, someone comes to oDesk, post a job, workers apply to that job. Predict which of these workers is most likely to be hired on that job. That was the narrow question. And so why is that a good question? Because we have a whole awesome set of tools now to solve that kind of a problem exactly. How do we do it? Take a lot of past data of past jobs, past applicants, past hires that were made. Then we ask these crazy big black box algorithms, "All right, do the best job you can predicting who's going to get hired on this job with these applicants." And we use that data to test how well these algorithms are doing. That's machine learning in 30 seconds basically. So we're working on this problem. Great.

And then I kind of poked my head up a little bit. I go, "Why are we working? What is this going to do?" Well, it turns out the reason these kinds of things are important is they get used to make decisions. So what kind of decision do you make with that? Well, one thing you do is you say, "Well, if I could predict who's most likely to be hired, then I should just rank people based on that, and that would be a good matching algorithm. That'd be a good way to sort and triage applicants for employers when they're screening, trying to figure out who to interview, who to hire." Great. Sounds pretty natural.

And then you think about it a little bit, and this to me, it's such a passion project to get people to understand that this is why the humans in the loop that help us in businesses and making sense of data are so critical, is the following problem. If you think about it a little bit, you realize what that algorithm is doing, it's really just picking up on patterns in past data. So yeah, that's great. This person is likely to be hired. But what we really want is something different. We're trying to add value by ranking people.

So to give another example that's similar to this, when you're a marketing manager, and you've got a cracked data science team that's built a long-term value, lifetime value model for you, you're not going to get in trouble with anyone if you send your highest value promotions to the highest LTV customers, right? Who's going to blame you for that? Because you're like, "This person is worth a lot, and I sent them this promotion." Say that in your monthly report, nobody's going to give you a hard time. But the problem with that way of thinking is actually predicting what their lifetime value is isn't really the question. The question is, how much more are they going to spend on my platform because I sent them that promotion? That's a very different thing. It's a differential rather than an absolute. I'm not interested in their absolute LTV. I'm absolutely interested in the difference in their LTV because I sent them this promotion.

And when you look at it that way, what you realize can happen is picking up on patterns because of good predictions, right? Finding the people that have high LTV because you predicted that is very different than making good decisions, which is about saying the difference in their LTV is going to be higher because I sent them this promotion. I love this example, because I taught a class here at Stanford. It was like an executive education class. We had all the executives from a company in the room, and one of the people in the room was the chief marketing officer. And I just asked this question like, "Hey, okay, let's say you got this great LTV model, who would you send the promotions to?" It's like, "Definitely the highest LTV people," and there's a CMO in the room. And so it's a little bit of a delicate situation, like pushing back a

little bit, right? I do want to be clear, there's reputational reasons you might do that anyway. I mean, I'm not trying to get away from that. But just to make the narrow point that predicting is about picking up patterns, but making decisions, it's about thinking about these differences.

Now, why is that important? Because we learn in high school, correlation is not causation. That's a phrase everybody has heard all over the place. What does that have to do with this? Well, when we teach people to build machine learning models, we're asking them to make predictions, we're asking them to find correlations. Prediction is inherently about correlation. But when we ask people to make decisions, we're asking them to think about causation. "If I make this decision, then will I actually increase the net value of my business? Will I have by sending the promotion, increased the likelihood that this person is going to spend more on my platform?" And so the first and most important thing that I feel very strongly about in what would I get a data scientist to do is no matter who they are, even if it was that person in the weeds thinking about building this prediction model for hiring, get them to be thinking in the back of their mind always that their goal is to help the business make decisions. And that the distinction between causation and correlation matters a lot. We can talk a lot more about how does that play out in terms of their day-to-day work. But at least at a starting point, you have to recognize that the first step is always recognition that prediction isn't the same thing as making decisions.

中文翻译：

在今天的许多公司里，你要求数据科学家做的主要工作就是构建所谓的“机器学习模型”。现在，“机器学习模型”对不同的人来说已经意味着很多不同的东西。我将专注于一些非常具体的东西：你要求他们预测某些事情。2012年我开始在oDesk工作时，我身上有一件有趣的事：我之所以去oDesk，是因为在那之前我有大约10年的学术生涯，一直在构建事物的数学模型。在那之前，我并不算是一个真正的数据科学家。我预想的是，我进入工业界，人们会告诉我：“嘿，看数据有多重要。”我的眼界确实大开了。

我被要求思考的第一件事是：好，有人来到oDesk发布了一个职位，工人们申请了这个职位。请预测这些工人中哪一个最有可能被录用。这是一个很窄的问题。为什么这是一个好问题？因为我们现在有一整套很棒的工具来精确解决这类问题。我们怎么做？获取大量过去职位、过去申请人、过去录用情况的数据。然后我们让这些疯狂庞大的黑盒算法去跑：“好，尽你所能预测在这些申请人中谁会被录用。”我们用这些数据来测试这些算法的表现。这就是30秒版的机器学习。所以我们正在研究这个问题。太棒了。

然后我稍微抬起头想了想：“我们为什么要研究这个？这有什么用？”事实证明，这类事情之所以重要，是因为它们被用来做决策。那么你用它做什么样的决策呢？你会说：“好吧，如果我能预测谁最有可能被录用，那么我就应该根据这个对人进行排名，这将是一个很好的匹配算法。这将是雇主在筛选、面试和录用时，对申请人进行排序和分流的好方法。”太棒了，听起来很自然。

但如果你再深入思考一下——这正是我非常热衷于让人们理解的一点，即为什么在业务中帮助我们理解数据的“人在回路”(Humans in the loop)如此关键——问题在于：如果你思考一下，你会意识到那个算法在做的，实际上只是在捕捉过去数据中的模式。所以，是的，这很好，这个人很可能被录用。但我们真正想要的是不同的东西。我们试图通过对人进行排名来增加价值。

举一个类似的例子：当你是一名营销经理，你拥有一支精锐的数据科学团队为你建立了一个长期价值(LTV，终身价值)模型。如果你把最高价值的促销活动发送给LTV最高的客户，没人会找你麻烦，对吧？谁会责怪你呢？因为你会说：“这个人价值很高，所以我给他发了促销。”在你的月度报告里这么说，没人会难为你。但这种思维方式的问题在于：预测他们的LTV并不是真正的问题。真正的问题是：因为我发送了那个促销，他们会在我的平台上多花多少钱？这是一个非常不同的东西。这是一个“增量”(Differential)而不是“绝对值”(Absolute)。我对他们的绝对LTV不感兴趣，我绝对感兴趣的是因为我发送了促销而导致的LTV差异。

当你这样看时，你会意识到：基于良好的预测来捕捉模式（比如因为你预测到了，所以找到了高LTV的人），与做出良好的决策（比如因为我发送了促销，他们的LTV增量会更高）是非常不同的。我喜欢这个例子，因为我在斯坦福教过一门课，那是一门高管教育课。房间里坐满了公司的所有高管，其中一位是首席营销官(CMO)。我问了这个问题：“嘿，假设你有了这个很棒的LTV模型，你会把促销发给谁？”回答是：“绝对是LTV最高的人。”而CMO就在房间里。所以这是一个有点微妙的情况，需要稍微反驳一下，对吧？我想澄清一点，出于

品牌声誉的原因，你可能还是会这么做，我不是要否定这一点。但只是为了说明一个狭义的观点：预测是关于捕捉模式，而决策是关于思考这些“差异”。

为什么这很重要？因为在高中就学过，“相关性不等于因果关系”。这句话每个人都听过。这和我们讨论的内容有什么关系？当我们教人们构建机器学习模型时，我们是在要求他们做预测，要求他们寻找相关性。预测本质上是关于相关性的。但当我们要求人们做决策时，我们是在要求他们思考因果关系。“如果我做了这个决策，我是否真的增加了业务的净价值？通过发送促销，我是否增加了这个人以后在平台上消费更多的可能性？”所以，对于“我会让数据科学家做什么”，我感到非常强烈的第一点也是最重要的一点是：无论他们是谁，即使是那个在基层思考构建招聘预测模型的人，也要让他们在脑海深处始终记住，他们的目标是帮助业务做决策。因果关系和相关性之间的区别至关重要。我们可以进一步讨论这在他们的日常工作中是如何体现的。但至少在起点，你必须认识到第一步永远是意识到：预测并不等同于做决策。

[00:35:27] [Lenny]

English:

So the takeaway here is as a data team and as a data scientist on the team, is help the business make predictions. Are there a couple more examples you could share of just what is an example of a decision that you think they often should be making and using data to help them with?

中文翻译：

所以这里的观点是，作为数据团队和团队中的数据科学家，要帮助业务做预测（注：此处 Lenny 顺着语境说成了预测，但 Ramesh 强调的是决策）。你还能分享几个例子吗？比如你认为他们通常应该做出哪些决策，并使用数据来提供帮助？

[00:35:41] [Ramesh Johari]

English:

Maybe the right frame of reference for this, and the word that an academic would use is causal inference. So what we're changing from is machine learning to causal inference. So let's think that through in a couple of different use cases that are related to that marketplace data science flywheel I talked about earlier. Finding matches, making matches, and then learning about matches. Finding matches, like you said, a core part of that is search and recommendation, and each of those relies on rankings. So I want to be able to rank order. Let's say I go do a search on Airbnb. On a rank order, the different listings in the marketplace, right? At some level, it's true that what I'm trying to do there is I'm trying to just predict, what are you going to like the most?

But I think there's an important piece of that also, which is that I want to think a little bit about the distinction between two different ranking algorithms. That's the real decision that's being made. And when I think about the distinction between two different ranking algorithms, I don't want to be only comparing them in terms of how well they recreate the choices people made in the past. The way I'm really going to evaluate those is in my market, does one of those lead to better matches or more matches than the other one, right?

So Airbnb as a business, what are the most obvious core metrics? It's bookings and revenue. So you're going to want to ask a very basic question. If I use the ranking algorithm Lenny just developed last night versus the ranking algorithm Ramesh developed last week, does Lenny's ranking algorithm lead to more bookings than Ramesh's ranking algorithm? And it's so important to put it that way starkly, because that's so different a question than, does Lenny's ranking algorithm do a better job of predicting over the

last two years what bookings people made than Ramesh's ranking algorithm? So that's I think at that level.

Then we talked a little bit about ranking at the point of making a match, and I think that's where this hiring issue popped up. Because in the end, while we might have these predictive algorithms to rank who you're going to hire, that's not the important question. Interestingly, the important question is actually to evaluate the quality of the match that's made. And we would do that through the next step of that flywheel. We'd ask ourselves, what ratings did they give back to that freelancer? Do they hire that freelancer again? So you're comparing two different algorithms not through their ability to recreate the past, but their ability to make matches in the future that can be objectively evaluated to say, "Hey, I increased the value of the business. I actually made better matches this way." And then rating systems, I think we could talk quite a bit about a similar phenomenon there too.

中文翻译:

也许对此最合适参考框架，或者学术界会用的词，是“因果推断”(Causal inference)。所以我们要从“机器学习”转向“因果推断”。让我们通过我之前提到的市场数据科学飞轮的几个不同用例来思考：寻找匹配、达成匹配、学习匹配。

寻找匹配，正如你所说，核心部分是搜索和推荐，而这两者都依赖于排名。所以我希望能进行排序。假设我在 Airbnb 上搜索，对市场中的不同房源进行排序。在某种程度上，我确实在尝试预测你最喜欢哪一个。但我觉得还有很重要的一点，那就是我要思考两种不同排名算法之间的区别。这才是真正要做的决策。当我思考两种不同排名算法的区别时，我不想仅仅比较它们在还原人们过去选择方面的表现。我真正评估它们的方式是：在我的市场中，其中一个算法是否比另一个带来了更好或更多的匹配？

对于 Airbnb 这样的业务，最明显的核心指标是什么？是预订量和收入。所以你会想问一个非常基本的问题：如果我使用 Lenny 昨晚刚开发的排名算法，对比 Ramesh 上周开发的排名算法，Lenny 的算法是否比 Ramesh 的算法带来了更多的预订？这样直白地提问非常重要，因为这与“Lenny 的算法在预测过去两年的预订情况方面是否比 Ramesh 的算法做得更好”是完全不同的问题。这是第一个层面。

然后我们聊到了达成匹配时的排名，这就是招聘问题出现的地方。因为最终，虽然我们可能有预测算法来对你可能雇佣的人进行排名，但这并不是最重要的问题。有趣的是，最重要的问题实际上是评估所达成的匹配的质量。我们会通过飞轮的下一步来做到这一点：我们会问，他们给那个自由职业者打了多少分？他们会再次雇佣那个自由职业者吗？所以，你比较两种不同算法，不是看它们重现过去的能力，而是看它们在未来达成匹配的能力，并且这种能力可以被客观评估，从而得出结论：“嘿，我增加了业务价值，我通过这种方式达成了更好的匹配。”至于评价系统，我认为我们也可以讨论类似的现象。

[00:38:12] [Lenny]

English:

This episode is brought to you by Eppo. Eppo is a next generation A/B testing and feature management platform built by alums of Airbnb and Snowflake for modern growth teams. Companies like Twitch, Miro, ClickUp, and DraftKings rely on Eppo to power their experiments. Experimentation is increasingly essential for driving growth and for understanding the performance of new features. And Eppo helps you increase experimentation velocity while unlocking rigorous deep analysis in a way that no other commercial tool does. When I was at Airbnb, one of the things that I loved most was our experimentation platform, where I could set up experiments easily, troubleshoot issues, and analyze performance all on my own. Eppo does all that and more, with advanced statistical methods that can help you shave weeks off experiment time, and accessible UI for diving deeper into performance, and out of the box reporting that helps you avoid annoying prolonged analytic cycles. Eppo also makes it easy for you to share experiment insights with your team, sparking new ideas for the A/B testing flywheel. Eppo powers experimentation across every use case, including product, growth, machine learning, monetization, and

email marketing. Check out Eppo at geteppo.com/lenny, and 10x your experiment velocity. That's geteppo.com/lenny.

中文翻译:

本集节目由 Eppo 赞助。Eppo 是由 Airbnb 和 Snowflake 的校友为现代增长团队构建的下一代 A/B 测试和功能管理平台。Twitch、Miro、ClickUp 和 DraftKings 等公司都依赖 Eppo 来驱动他们的实验。实验对于驱动增长和了解新功能表现越来越重要。Eppo 能够帮助你提高实验速度，同时以其他商业工具无法做到的方式解锁严谨的深度分析。当我在 Airbnb 时，我最喜欢的事情之一就是我们的实验平台，在那里我可以轻松设置实验、排除故障并独立分析表现。Eppo 具备所有这些功能甚至更多，它拥有先进的统计方法，可以帮你缩短数周的实验时间，还有易于操作的 UI 用于深入研究表现，以及开箱即用的报告，帮你避免烦人且漫长的分析周期。Eppo 还让你能轻松与团队分享实验洞察，为 A/B 测试飞轮激发新创意。Eppo 支持各种场景下的实验，包括产品、增长、机器学习、盈利化和邮件营销。在 geteppo.com/lenny 查看 Eppo，让你的实验速度提升 10 倍。地址是 geteppo.com/lenny。

[00:39:29] [Lenny]

English:

Yeah, I would actually love to talk about rating systems, but there's an implication in everything you're describing of running an experiment versus looking at what would've happened in the previous world made. You've made change, run an experiment, see if it actually makes an impact on bookings and revenue. And that leads me to a question I wanted to ask, which is with experiments, there's kind of this classic challenge, and always elephant in the room of if you just run a bunch of experiments, you're kind of going to micro optimize, lead to these local maxima, and you may miss big opportunities and big unlocks if you're just extremely experiment driven. You spend a lot of time thinking about experimentation. What have you learned or what advice do you have for people to either be less worried about optimizing and missing something big, or just finding a balance with running experiments, but also creating opportunity to find a huge new opportunity?

中文翻译:

是的，我很想聊聊评价系统，但在你描述的所有内容中，都隐含着“运行实验”与“观察过去发生的事情”之间的对比。你做出了改变，运行实验，看看它是否真的对预订量和收入产生了影响。这引出了我想问的一个问题：关于实验，有一个经典的挑战，也是一个避不开的话题，那就是如果你只是运行一堆实验，你可能会陷入“微观优化”，导致“局部最优解”（Local maxima），如果你极端依赖实验驱动，你可能会错过巨大的机会和重大的突破。你花了很多时间思考实验。你学到了什么，或者你对人们有什么建议，让他们既不用太担心过度优化而错过大事，又能平衡好运行实验与寻找巨大新机会之间的关系？

[00:40:21] [Ramesh Johari]

English:

Yeah. First of all, I'm really glad you broached the E word. I was dancing around it, and I'm really glad that we talked about experiments. Because yeah, one of the big lessons of this recent conversation we've been having is just, how could you possibly know that difference without doing something like experimenting? So yeah, I am a big believer in experiments. I mean, I'll just lay those cards on the table. I love working with businesses that think experiments are important to helping make good decisions. Now all that said, I am also someone who feels pretty strongly about this exact issue that you're raising. It's just, you can't experiment your way out of everything.

And one frame I like to give people is that although you might say you're an experiment driven business, some businesses will proclaim, "We literally test everything." What that kind of leaves aside a little bit is there's a lot of degrees of freedom in what it means to test everything. Because ultimately, what's getting built and tested are choices that are made through the organizational structure, the data scientists, the PMs, the engineers, everybody's on... Before we're running experiments, we're actually thinking about even what's worth experimenting, what designs are we coming with? So that's one. And the other big one is, how long do we run these experiments? Okay, that's a big choice. And what I generally believe, and I think there's a paper we can link to later that I'll point your readers to as well that... Not my paper, from some folks at Microsoft.

What I generally believe is we're risk averse on both these two dimensions, that what people decide to test in a world that has promoted experimentation for everything tends to be more incremental by design. Okay? And we'll come back to that actually, answer the because in a second. So that's one. And two is people tend to run experiments for a long time, and probably longer than they should. Now, what do I mean by these two things? So what's interesting to me about this dynamic is experiments don't live in a vacuum. Companies have incentives. And in companies that really go all in on experimentation, one of the things that gets wrapped up in that is the incentives around experiments. Because if you go all in on experiments, a common thing you'll see is data scientists get judged based on how many wins they had that quarter. How do you get more wins?

Well, it's easier to get wins when you're being incremental. And because it's important to have wins, you have to run them long enough to demonstrate that they're really wins. You're less willing to cut something off in exchange for trying something riskier. So the big lesson of this Microsoft paper, it's called A/B Testing with what's called Fat Tails, which in lay terms just means you're running a business where there's potentially big opportunities out there if you look at the effects of the experiments that you run. There's a couple of lessons there about both trying a lot more stuff that's not all risk averse, and not necessarily running everything for so long. So really getting velocity up.

So you could see that there's a big incentive problem there, because the culture that says it's okay to fail big actually requires changing the terminology of wins. This is one of the things I hate most in A/B testing, I have to say. I get where it comes from. Experimentation was never historically in science about winners and losers. It'd be weird if it Ronald Fisher who's kind of the father of experimentation with his agriculture experiments talked about winners. I don't think that's necessarily how he talked about things. Experimentation is always very hypothesis driven. It's about, what are you learning?

And that's really an important distinction because what it means is if I go with something big, risky, and it, "fails," meaning that doesn't win. Nevertheless, if I was being rigorous about what hypotheses that's testing about my business, I'm potentially learning a lot. So a great example of this kind of thing is that there's an important feature of marketplaces is badging. So sometimes, it's really important to have badges on your top-rated profiles or whatever, when people are searching. And without going too far into the details, a common finding with badges is that badges you think are going to be great actually turn out to be terrible. And one reason they're terrible is they focus too much attention on the badged folks, and pull too much attention away from the unbadged folks. And if we judge that only in terms of winners and losers, you throw the baby out with the bath water, you're like, "Well that badging idea was terrible. So ditch that, no badges." But that's not what it's telling you. It's teaching you something about how inventory is being reallocated, how attention's being redirected through the badges. And you really want to think not in terms of winning and losing, but learning.

So learning is a win. And I feel that that's a cultural thing fundamentally. It's very hard to somehow attach dollars and cents at the top to data scientists running experiments that fail, but learn. And ultimately, I think getting into that space where you experiment more, meaning you don't run all your experiments for quite as long and you accept the willingness to try experiments that are into the tails where you might fail

bigger is a cultural thing. It's about saying that, "We're allowing that to be part of our social contract with our data scientists," or actually our employee contract with our data scientists, that not everything is just about how many launches you had and how many wins there were. It's okay to say, "That's how I want to use experimentation," but if you're going to use it that way, then I would say don't be a, "We experiment everything," business. Because then I think you need some other way to deal with these big changes that teach the whole company a lot, but maybe can't fall into the incentives you've created for your data scientists.

中文翻译:

是的。首先，我很高兴你提到了“实验”（Experiment）这个词。我刚才一直在绕圈子，很高兴我们聊到了实验。因为我们最近谈话的一个重要教训就是：如果不通过实验，你根本不可能知道那种差异。所以，是的，我是实验的坚定信徒。我把话挑明了：我喜欢与那些认为实验对辅助决策至关重要的企业合作。话虽如此，我也对你提出的问题深有感触。那就是：你不能靠实验解决所有问题。

我喜欢给人们提供的一个框架是：虽然你可能说自己是一家实验驱动的企业，有些公司甚至宣称“我们测试一切”。但这忽略了一点，即“测试一切”意味着在很多维度上都有自由度。因为最终，被构建和测试的东西是由组织结构、数据科学家、产品经理、工程师共同做出的选择。在运行实验之前，我们实际上是在思考什么值得实验，我们要拿出什么样的设计。这是第一点。另一个大问题是：我们要运行这些实验多久？这是一个重大选择。我普遍认为——稍后我们可以链接一篇微软研究人员写的论文，不是我写的——我普遍认为我们在两个维度上都是风险厌恶的。

第一，在推崇“万物皆可测试”的世界里，人们决定测试的东西往往在设计上更倾向于“增量式”的改进。第二，人们倾向于把实验运行很长时间，可能比实际需要的还要长。这两点是什么意思？对我来说，这种动态有趣的地方在于实验并不是在真空中进行的。公司有激励机制。在那些全力投入实验的公司里，实验周围的激励机制也会卷入其中。因为如果你全力投入实验，你会看到一个普遍现象：数据科学家是根据他们那个季度获得了多少次“胜利”（Wins）来考核的。你如何获得更多胜利？

显然，当你做增量式改进时，更容易获得胜利。而且因为获得胜利很重要，你必须让实验运行足够长的时间，以证明它们真的是胜利。你不太愿意为了尝试风险更大的东西而中断当前的实验。所以那篇微软论文（题目叫《带有肥尾效应的 A/B 测试》，Fat Tails）的核心教训是：如果你观察实验的效果，你会发现业务中可能潜藏着巨大的机会。那里有几个教训：一是尝试更多非风险厌恶的东西，二是不必把所有实验都跑那么久。也就是真正提高实验速度。

所以你可以看到这里有一个巨大的激励问题。因为一种认为“大失败也没关系”的文化，实际上需要改变对“胜利”的定义。我必须说，这是我在 A/B 测试中最讨厌的事情之一。我明白它的来源，但在科学史上，实验从来不是关于赢家和输家的。如果实验之父罗纳德·费舍尔（Ronald Fisher）在做农业实验时谈论“赢家”，那会很奇怪。我不认为那是他谈论事情的方式。实验总是由假设驱动的。它是关于：你学到了什么？

这是一个非常重要的区别。这意味着如果我尝试了一些重大的、有风险的东西，结果它“失败”了（即没有胜出），但如果我严谨地对待这个关于业务的假设，我可能会学到很多东西。一个很好的例子是市场中的“徽章”（Badging）功能。有时，在搜索结果中给顶级个人资料打上徽章非常重要。在不深入细节的情况下，关于徽章的一个常见发现是：你认为会很棒的徽章实际上表现很糟糕。原因之一是它们把过多的注意力集中在有徽章的人身上，而从没徽章的人那里夺走了太多的注意力。如果我们只用赢家和输家来衡量，你就会把孩子连同洗澡水一起倒掉，你会说：“徽章这个主意太烂了，取消它，不要徽章。”但这并不是实验告诉你的。它是在教你关于库存如何被重新分配、注意力如何通过徽章被重新引导的知识。你真正应该思考的不是输赢，而是学习。

所以，学习就是一种胜利。我觉得这从根本上是一种文化。很难将金钱利益与数据科学家运行的失败但有收获的实验挂钩。最终，我认为进入一个能进行更多实验的状态——意味着不把所有实验都跑那么久，并愿意尝试那些可能导致更大失败的“长尾”实验——这是一种文化。这关乎于说：“我们允许这成为我们与数据科学家之间社会契约（或者说雇佣合同）的一部分”，即并不是所有事情都只关乎你发布了多少功能、获得了多少次胜利。你可以说“这就是我想使用实验的方式”，但如果你真的这么做，我就建议你不要成为一家“测试一切”的公司。因为那样的话，你需要另一种方式来处理那些能让全公司学到很多、但可能无法纳入你为数据科学家设定的激励机制中的重大变革。

[00:46:04] [Lenny]

English:

This badging example is, I don't know if you're referring to the Airbnb example, but I actually led the launch of Superhost at Airbnb, which is the ultimate badge on Airbnb. And there was a lot of concern from the data team that it would destroy the marketplace, because they've built, as you've described, this very well-crafted ranking algorithm, with just a prediction of exactly as you described, which listings that guest is most likely to book and be successful booking. And then we're about to throw a badge on random listings in the results. And so this one data scientist on our team's like, "No, we can't do this. This is insane. We're going to destroy it all." And we still went ahead with it. We ran an experiment showing the badge to some people and some not, actually, it was no impact at all. Which Superhost itself had no impact at all on the business as far as we could tell initially, which is also bittersweet because it felt like, "Why did we even work on this thing?" There was a slight benefit where a host felt better, they felt more satisfied with being a host, but I went exactly through what you described, so that's pretty funny.

中文翻译:

这个徽章的例子，我不知道你是不是在指 Airbnb 的例子，但我实际上领导了 Airbnb “超赞房东” (Superhost) 的发布，那是 Airbnb 上终极的徽章。当时数据团队非常担心这会摧毁市场，因为正如你所描述的，他们构建了一个非常精妙的排名算法，精确预测哪些房源是房客最有可能预订并成功入住的。然后我们要在搜索结果中给一些房源打上徽章。我们团队的一位数据科学家当时说：“不，我们不能这么做。这太疯狂了。我们会毁掉一切的。”我们还是推进了。我们运行了一个实验，给一部分人显示徽章，另一部分人不显示。实际上，最初对业务完全没有影响。据我们最初所知，超赞房东本身对业务没有任何影响，这让人感到苦乐参半，因为你会觉得：“我们为什么要费劲做这个？”虽然有一点点好处，比如房东感觉更好了，他们对作为房东感到更满意，但我经历的过程和你描述的完全一样，这挺有意思的。

[00:47:11] [Ramesh Johari]

English:

Without necessarily going into the weeds on the data science of Superhost, I think there's a lot wrapped up in what you said. I guess another thing I'll say is that I'm a big believer that you don't throw your understanding of the business out the window when you process experiment results. And it's partly, I guess what I mean by this is data science is really about accumulation of evidence. It's never about one finding in isolation. And so another kind of trap I think is to sometimes say, "Well, I hit stat sig on my A/B test, green light. It's all go." And I think you had Ronny Kohavi on your show, and he made a similar point that there are different levels of evidence, and just having an outlier A/B test that goes against everything you believe about your business doesn't mean that you somehow have controverted all your knowledge. And I think that's one side of it.

The other thing is you can't always measure everything that's important that's needed to really develop a full sense. So with Superhost, one of the things that's hard to measure is the long-term impact of Superhost. Because in the short run, Superhost causes a rebalancing of inventory. There's going to be winners and losers. Part of Superhost is actually about retaining hosts that get the badge over a longer period of time. Recognizing that hypothesis actually says something about maybe how long the experiment needs to be run or what kinds of data analyses need to be done. And in the end, if you can't do that, you can't run it long enough, or you can't do that data analysis due to sparsity of data or lack of data to address the question, it matters what you bring to the table. What are your beliefs about that?

So what I like to tell people to do there is I like to push people to be what's called quantified rather than data-driven, which is, okay fine, some things we can't measure. But maybe you've got a leadership team

with different beliefs about what they think the retention value of Superhost is going to be, and they might be all over the place. You can process your experiment results in the context of these competing beliefs. It's almost like a prediction market kind of a thing. And start asking, "Well okay, if this is what we believe about our business, this is what the data is telling us out of the experiment, let's put those two together and ask, is this enough for us to make the bet that we're still going to go with it?" Even though maybe that short-term test you ran was flat.

中文翻译:

在不深入探讨超赞房东数据科学细节的情况下，我认为你刚才说的话里包含了很多信息。我想说的另一件事是：我坚信在处理实验结果时，你不能把对业务的理解抛诸脑后。我的意思是，数据科学本质上是证据的积累，而不是孤立地看某一个发现。我认为另一个陷阱是，有时人们会说：“好吧，我的 A/B 测试达到了统计显著性 (Stat sig)，绿灯，全速前进。”我想你之前请过 Ronny Kohavi 上节目，他也表达过类似的观点：证据有不同的层级，仅仅一个与你对业务的所有认知相悖的异常 A/B 测试结果，并不意味着你所有的知识都被推翻了。这是其中一方面。

另一方面是，你并不总能衡量出开发全方位认知所需的所有重要因素。对于超赞房东，很难衡量的一件事是它的长期影响。因为在短期内，超赞房东会导致库存的重新平衡，会有赢家和输家。超赞房东的一部分意义实际上在于长期留住那些获得徽章的房东。意识到这个假设，实际上决定了实验需要运行多久，或者需要进行什么样的数据分析。最终，如果你做不到这一点——无法运行足够长的时间，或者由于数据稀疏或缺乏数据而无法进行分析——那么你带到谈判桌上的东西就很重要了。你对此的信念是什么？

所以我喜欢告诉人们的是，我倾向于推动人们变得“量化” (Quantified) 而不是单纯的“数据驱动” (Data-driven)。也就是说，好吧，有些东西我们无法衡量，但也许你的领导团队对超赞房东的留存价值有不同的信念，这些信念可能各不相同。你可以在这些相互竞争的信念背景下处理你的实验结果。这有点像预测市场。然后开始问：“好吧，如果这是对我们对业务的信念，而这是实验告诉我们的数据，让我们把这两者结合起来问：这是否足以让我们下注继续推进？”即使你运行的短期测试结果是持平的。

[00:49:59] [Lenny]

English:

That's actually exactly how I think of Superhost looking back. It was a great idea. I'm really happy. I can't even imagine Airbnb without that, even though there's no evidence, at least initially, that it made any impact. I'm guessing they looked at it again, and maybe there's something that came out of it. But even if it had no impact, it just feels like it made the marketplace better. And that was a big learning for me. It doesn't need to always drive a metric that you can measure. There's just like, this is the way it should work.

中文翻译:

回想起来，我对超赞房东的看法正是如此。这是一个伟大的主意。我很高兴做了这个，我甚至无法想象没有它的 Airbnb 会是什么样，尽管至少在最初，没有任何证据表明它产生了影响。我猜他们后来又重新审视了，也许发现了一些东西。但即使它没有影响，它也让人觉得让市场变得更好了。这对我来说是一个很大的学习：它并不总是需要驱动一个你能衡量的指标。有时只是觉得“这就是它应该运作的方式”。

[00:50:23] [Ramesh Johari]

English:

So one of the reasons the thing you said happens is because marketplaces are a little bit like a game of whac-a-mole, okay? And what I mean by that is, so narrowly in the context of Superhost, because you're redirecting attention to some hosts at the expense of... It's not even obvious if bookings can really go up.

Maybe you get lucky and maybe you get a bunch more bookings. One reason you probably wouldn't expect that in the first place is there's only a limited number of Superhosts. How many more bookings are they going to be absorbing because of all this extra attention? And you're taking attention away from other people. Without doing any data analysis, my prior would've been that booking should probably go down.

And one example that I came across with one of the companies I worked with that I love is we were working together over a period of time, and in a month, we looked at some of the data and it suggested that our new supply side was having a pretty bad experience. Say, "We got to do something about this." So what we decided to do is build some custom bespoke features that were really going to direct them to more experienced folks on the other side of the market. Good. And then lo and behold, pretty soon those metrics start to look better. But then we're looking at it, we're like, "Wait a second. Now the existing folks on the other side are having a worse experience."

So you kind of whiplash around. You're like, "Wait a second, we better do something about that." So we take them, we try to match them up with the more experienced folks. And now suddenly a month after that you're like, "Wait a second." And your metrics just keep moving around. And that's because the whac-a-mole game here is ultimately, a lot of marketplace management is moving attention and inventory around. Sometimes you get lucky and you really expand the pie for everybody. But I think Servaes Tholen, who was CFO at Upwork that I got to know there and then went to Thumbtack later, he had this line when he came to visit our class that I love, which is, "You have to recognize when you run marketplaces that many of the changes that are most consequential create winners and losers. And rolling with those changes is about recognizing whether the winners you've created are more important to your business view than the losers you've created in the process." And it's a hard reality, because nobody likes to articulate the idea that a feature change is hurting some of the people in your marketplace. But because of this fundamental constraint baked into how marketplaces work, many of the things that we would choose to do and the reallocation they create can't necessarily create observed high expanding wins in the short run. You're often making bets that that's where you're headed, partly through the reallocation that you're doing right now. And so I think that's interesting about Superhost to me is that partly points to thinking about, what's the objective you would've defined, the metric you would've defined in the short run that captures this idea of a trade-off?

中文翻译:

你所说的情况之所以发生，原因之一是双边市场有点像打地鼠游戏。我的意思是，具体到超赞房东的背景下，因为你把注意力重新导向了一些房东，而代价是……预订量是否真的能上升甚至都不明显。也许你运气好，预订量增加了一堆。但你最初不应该对此抱有期望的一个原因是：超赞房东的数量是有限的。由于这些额外的关注，他们到底能多消化多少预订？而且你正在从其他人那里夺走关注。在不做任何数据分析的情况下，我的“先验”（Prior）判断是预订量可能会下降。

我合作过的一家公司有一个我很喜欢的例子：我们合作了一段时间，某个月我们看数据，发现我们的新供应端体验非常糟糕。我们说：“得做点什么。”于是我们决定开发一些定制功能，将他们引导至市场另一端更有经验的用户。结果，很快这些指标就开始好转了。但接着我们观察发现：“等一下，现在另一端原有的资深用户体验变差了。”

于是你就像遭遇了挥鞭效应一样，心想：“等等，我们得处理一下。”于是我们又去调整，尝试把他们与更有经验的人匹配。结果一个月后，你又会惊呼：“等等，”你的指标就这样不停地变动。这是因为打地鼠游戏的本质在于，双边市场的管理最终其实是在调配注意力和库存。有时你运气好，真的为每个人都做大了蛋糕。但我认识的 Upwork 前 CFO（后来去了 Thumbtack）Servaes Tholen 来我们班上讲课时说了一句我很喜欢的话：“在经营市场时，你必须意识到，许多影响最深远的变革都会产生赢家和输家。而应对这些变化的关键在于，你要意识到你创造的赢家是否比你在这个过程中产生的输家对你的业务愿景更重要。”这是一个残酷的现实，因为没人喜欢明确表达“某个功能改动正在伤害市场中的一部分人”。但由于市场运作方式中固有的这种根本约

束，我们选择做的许多事情以及它们带来的重新分配，在短期内未必能产生可观察到的、大幅扩张的胜利。你通常是在下注，赌那是你前进的方向，而这部分是通过你现在正在做的重新分配来实现的。所以我觉得超赞房东有趣的地方在于，它在某种程度上指向了一个思考：你在短期内会定义什么样的目标或指标，来捕捉这种“权衡”（Trade-off）的想法？

[00:52:42] [Lenny]

English:

That's a great way to think about it. I wanted to come back to this idea you're sharing of maybe you should run experiments more quickly, not wait for stat sig, have a culture of learning versus impact. In practice, it's very difficult, because people are measured by impact. There's performance reviews, there's promotions, there's how much impact did this team drive, are going to look at their experiment results? You've worked with a lot of marketplace companies, a lot of different companies. Is there anything you've seen about something you could do to help the company shift and actually work this way, while also recognizing success, and who's doing great, who's not, which team's driving impact, who's not?

中文翻译:

这是一个很好的思考方式。我想回到你分享的那个观点：也许应该更快地运行实验，而不是等待统计显著性，建立一种“学习导向”而非“影响导向”的文化。在实践中这非常困难，因为人们是根据“影响”（Impact）来衡量的。有绩效评估、有晋升，大家会看这个团队带来了多少影响，会看他们的实验结果。你与很多市场公司和不同类型的公司合作过。你有没有见过什么方法可以帮助公司转型并真正以这种方式运作，同时又能识别成功，看谁做得好，谁做得不好，哪个团队带来了影响，哪个没有？

[00:53:22] [Ramesh Johari]

English:

Interestingly, it's actually an active area of research for me now. What I mean by active area of research is I care a lot about the incentives that we create for data science through how we set up reward mechanisms. So there's a couple things I think that could be helpful, that are maybe there may be a little bit less about... Maybe I'm not going to directly answer the question you ask, because I think that's a hard one, right? I think I recognize that measurement on impact is critical. Well, let me answer that actually from the most obvious way first. I think there's a cultural issue here that's really critical. One of the things I often find is that my PhD students, our PhD students here often go off and get great data scientist jobs. And in one sense, they're doing amazing stuff. They apply really technically sophisticated methods. But when I look at the problems they're working on, they're often more at the margins of the business than they should be.

And it's a cultural thing. It's basically because if you're measured narrowly on impact and that's all anyone sees around you, then it's very hard to engage with the creative aspect of business change and the strategic aspects of business change. So the cultural aspect there is, I think it's partly incumbent on the leaders to expect something more of their data scientists. And what I mean by expect more is that you expect them to do more than deliver narrowly defined, statistically rigorous results to you in their reports. You're actually expecting them to talk also about what they're learning about the business in the process. So where that's headed is there's this concept of being hypothesis driven, which is like the technical phrase. What does that mean? Again, in a more lay sense.

What it means is tests aren't going to be defined only in terms of winners and losers, that each test should also say something about what will we learn about a business flow, a funnel, preferences of the guests, preferences of the hosts. What will we learn about their demand elasticity if we're changing prices

around? These kinds of things. So it's possible to articulate in an experiment doc, a launch doc, what are the hypotheses that are being tested? So that's one thing I would say is just culturally, setting the norms that learning is part of the discourse, and it's expected actually I think is important.

But the other thing I would say that's maybe a little bit more about programmatically, what could a data science platform team do? A funny thing about experiments is that we throw past learning away effectively. And this is just an artifact of how we analyze experiments, that the statistical methods used typically, P-values, confidence intervals, these fall into a branch of statistics known as frequentist statistics. And the idea behind frequentist statistics without being overly technical is just I let the data speak for itself. There's no beliefs brought to the table about where that data came from. But if you think about this in a company, in A/B testing a company, it's a weird thing, right? Because I might've run 1,000 A/B tests in the past on this exact same button, or call to action, or color, and now I am going to completely ignore that and focus only on this.

So there's ways to take the past into account, to build what's called a prior belief before I run an experiment, and now take the data from the experiment, connect it with the prior, to come up with a conclusion of, "Okay, in light of the past plus this experiment, what's it telling me about the future?" And that falls broadly under the category of what's called Bayesian A/B testing. So that's one of the things I think can help culturally, weirdly. It's a super technical thing, but I think it can help culturally, because what it's doing is it's now rewarding people for contributing information to that prior. And I think it then becomes possible to say, "Your experiment that failed actually moved our prior." And that's an important thing, because by doing so, you're now altering how we're going to think about this flow or this pricing plan in all future experiments. So there's an information positive externality, positive network effect that's generated for the rest of your business if I can somehow encode what you learned into the analysis of future experiments. So this is one thing. There's strong connection between the culture and incentives of A/B testing and the ability to actually incorporate past learning into these prior beliefs.

中文翻译:

有趣的是，这实际上是我现在的一个活跃研究领域。所谓活跃研究领域，是指我非常关注我们通过设置奖励机制为数据科学创造的激励。我认为有几件事可能会有帮助，虽然它们可能不那么……也许我不会直接回答你问的那个问题，因为我觉得那很难，对吧？我承认衡量影响是至关重要的。好吧，让我先从最显而易见的方式来回答。我认为这里有一个非常关键的文化问题。我经常发现，我的博士生们毕业后去做了很棒的数据科学家工作。从某种意义上说，他们做着了不起的事情，应用了技术上非常复杂的方法。但当我审视他们正在研究的问题时，这些问题往往比应有的更处于业务的边缘。

这是一种文化问题。基本上是因为如果你被狭隘地根据“影响”来衡量，而且周围的人只看这个，那么你就很难参与到业务变革的创意层面和战略层面。所以文化层面在于，我认为领导者有责任对他们的数据科学家有更高的期望。我所说的“更高期望”是指，你期望他们做的不仅仅是在报告中向你交付定义狭窄、统计严谨的结果。你实际上还期望他们谈谈在这个过程中他们学到了什么关于业务的知识。这就是所谓的“假设驱动”(Hypothesis driven)，这是一个技术词汇。通俗地说，这意味着什么？

这意味着测试将不再仅仅根据赢家和输家来定义，每个测试还应该说明：关于业务流程、漏斗、房客偏好、房东偏好，我们学到了什么？如果我们调整价格，关于他们的需求弹性，我们学到了什么？诸如此类。在实验文档或发布文档中明确表达正在测试哪些假设是可行的。所以我要说的一点是，从文化上设定“学习是对话的一部分”且是被期待的规范，我认为这很重要。

另一件事可能更偏向程序化，即数据科学平台团队能做什么？实验中一个滑稽的现象是，我们实际上把过去的学习成果扔掉了。这只是我们分析实验方式的一个产物，通常使用的统计方法（如 P 值、置信区间）属于统计学的一个分支，称为“频率派统计学”(Frequentist statistics)。频率派统计学的核心思想（不涉及太多技术细节）就是让数据自己说话，不带入任何关于数据来源的信念。但如果你在一家公司里思考 A/B 测试，这很奇怪，对吧？因为我过去可能针对同一个按钮、同一个行动号召 (CTA) 或颜色运行过 1000 次 A/B 测试，而现在我要完全忽略这些，只关注这一次。

所以，有一些方法可以将过去考虑在内，在运行实验之前建立所谓的“先验信念”（Prior belief），然后获取实验数据，将其与先验结合，得出一个结论：“好吧，结合过去和这次实验，它告诉了我们关于未来的什么？”这广义上属于“贝叶斯 A/B 测试”（Bayesian A/B testing）的范畴。奇怪的是，我认为这在文化上会有所帮助。这是一个超级技术性的东西，但它能提供文化上的帮助，因为它现在奖励人们为那个“先验”贡献信息。我认为这样一来，就有可能说：“你那个失败的实验实际上改变了我们的先验。”这是一件大事，因为通过这样做，你改变了我们在未来所有实验中对这个流程或定价计划的思考方式。如果我能以某种方式将你学到的东西编码到未来实验的分析中，就会为你的其他业务产生一种信息的正外部性或正网络效应。所以这是其中一点：A/B 测试的文化和激励机制，与将过去学习成果纳入先验信念的能力之间有着强有力联系。

[00:57:35] [Lenny]

English:

I love that you're doing research in this area. We should bring you back when you've completed it and have the ultimate answer for everyone to change how they operate.

中文翻译:

我很喜欢你正在这个领域做研究。等你完成了，我们应该再请你回来，给每个人一个改变运作方式的终极答案。

[00:57:42] [Ramesh Johari]

English:

Yeah, one of the great things about professors is we never complete anything and never have ultimate answers.

中文翻译:

是啊，教授的一大特点就是我们从不完成任何事情，也从没有终极答案。

[00:57:47] [Lenny]

English:

Oh boy.

中文翻译:

噢，天哪。

[00:57:47] [Ramesh Johari]

English:

Yeah, I'll do my best though.

中文翻译:

不过我会尽力的。

[00:57:50] [Lenny]

English:

This touches on a really interesting concept that you shared with me around how, just learning isn't free. People think that they could just learn a bunch of stuff and there's not a cost to it. I'd love for you to just chat a bit about what that means.

中文翻译:

这触及了你曾跟我分享过的一个非常有趣的概念：学习并不是免费的。人们认为他们可以随便学到一堆东西而不需要付出代价。我很想听你聊聊这意味着什么。

[00:58:02] [Ramesh Johari]

English:

Let me start with an anecdote, that I just absolutely love this anecdote. I use it every year in class. So I was talking to a real estate platform, and they had a marketing data science manager who's basically responsible, as many marketing managers are, for allocation of ad spend across different channels. And what they discovered had happened at the end of the year is in one hand, the team had done great, but the manager had held out some subset of arriving visitors, not showing them any of the innovations they were making.

中文翻译:

让我从一个轶事开始，我绝对爱死这个故事了，每年上课我都会讲。我曾与一家房地产平台交流，他们有一位营销数据科学经理，像许多营销经理一样，负责在不同渠道分配广告支出。年底时他们发现，一方面团队表现出色，但另一方面，这位经理预留了一小部分访客作为“预留组”（Holdout group），不向他们展示团队正在做的任何创新。

[00:58:39] [Lenny]

English:

Like a holdout group?

中文翻译:

就像对照组/预留组那样？

[00:58:40] [Ramesh Johari]

English:

Yeah, exactly. What's called a holdout group in experimentation. And one thing about this holdout is it wasn't authorized. That's not the way things are supposed to work. They've got their ad spend, allocate out your ad spend, great. So at the end of the year, they looked at the hole out and they're like, "Wow, that cost us a couple million dollars, something in that range, and it's not a trivial amount of money. What's the deal? What were you thinking?" Basically. And of course the answer was, "Well, I get that I cost you that much, but number one, now you know what my team's worth. And number two, you would never have had that answer unless I'd done that on my own."

Now, why is that so powerful? I think what I find so interesting about experiments is that when you don't know something, it seems not even a question that you would allocate some of your samples to all options, right? Treatment and control. I have two different ways of doing something. I don't know which

one's better, so of course I'll give some samples to each. After the fact you're like, "Treatment was better. What the heck were we thinking? Why'd we give all those samples to control? That doesn't make any sense now." There's this great Seinfeld clip where he mentions getting a bill at the end of a large luxurious meal, and people stare at the bill like, "We're not hungry now. Why'd we order all this food?" So it's the same thing. I mean, you know treatment's better now. Why'd you waste all those samples on control?

And I think that is such a powerful observation that you have to put yourself in the frame of reference of when you didn't have the answer. And at that moment, what you're essentially saying to yourself is that it's worth paying to learn the answer. I think it sounds obvious the way we're saying it now, or this anecdote of the marketing manager and the holdout sounds obvious. What's culturally not baked in I think is that idea. And the reason I say it's not culturally baked in, by the way, is because of the language of winners and losers. Because if we use that language, we're implicitly saying is that we wasted time when we ran an A/B test on loser. If I reward you for shipping winners, then what I'm really telling you is all the time that you spent testing out failures was wasted time. And I think, of course, you don't want to keep data scientists around who regularly are just generating failures. That's not my point. But my point is there's a disconnect there. On one hand, we can all look at the story of this marketing manager and chuckle at it. And yet, every day we're instantiating language and processes that are reinforcing that same theme, which is essentially trying to say to you, "If you're wasting samples on things that don't ultimately end up being a winner, then the act of doing so is a failure." So I really feel that that idea that you have to pay to learn, again, it's a cultural thing, but it's also an education issue for businesses are populated by people of all stripes. Not everybody comes from a data science or experimentation background. And this idea that learning is costly is not natural, actually. It's not natural as a matter of human nature. It's certainly not natural as a matter of running a business.

中文翻译:

是的，没错。实验中所谓的“预留组”。关于这个预留组的一点是，它未经授权。事情本不该这样运作。他们有广告预算，你就该分配出去。所以年底时，公司看了看那个预留组，心想：“哇，这让我们损失了几百万美元，这可不是一笔小数目。怎么回事？你怎么想的？”基本上就是这样。当然，回答是：“好吧，我知道我让公司损失了那么多钱，但第一，现在你们知道我团队的价值了；第二，除非我自作主张这么做，否则你们永远不会得到这个答案。”

为什么这如此有力？我觉得实验有趣的地方在于：当你不知道某件事时，把样本分配给所有选项（实验组和对照组）似乎是理所当然的，对吧？我有两种做事方式，不知道哪种更好，所以当然各分一些样本。事后你会说：“实验组更好。我们到底在想什么？为什么要把那么多样本给对照组？现在看来这毫无意义。”《宋飞正传》（Seinfeld）里有一个精彩片段，他提到在大餐结束拿到账单时，人们盯着账单想：“我们现在不饿了，为什么要点这么多菜？”这是一回事。我是说，你现在知道实验组更好了，为什么还要在对照组上浪费样本？

我认为这是一个非常强有力地观察：你必须把自己置于“还没有答案”的参考框架中。在那一刻，你本质上是在对自己说，为了学到答案，付出代价是值得的。我们现在说起来听起来很显而易见，或者那个营销经理和预留组的故事听起来很显而易见。但我认为这种想法在文化上并没有根深蒂固。顺便说一下，我认为它在文化上没根深蒂固的原因在于“赢家和输家”的语言。因为如果我们使用这种语言，我们隐含的意思就是：当我们在“输家”上运行A/B测试时，我们浪费了时间。如果我因为你发布了“赢家”而奖励你，那么我实际上是在告诉你，你花在测试“失败”上的所有时间都是浪费。当然，你不想留住那些经常只产生失败的数据科学家，这不是我的重点。我的重点是这里存在脱节。一方面，我们都可以看着这个营销经理的故事付之一笑；但另一方面，我们每天都在实例化那些强化同一主题的语言和流程，本质上是在对你说：“如果你把样本浪费在最终没能成为赢家的事情上，那么这种行为本身就是失败。”所以我真的觉得“必须为学习付费”这个想法，再次强调，它是一个文化问题，也是一个教育问题，因为企业由各色人等组成，并不是每个人都有数据科学或实验背景。事实上，“学习是有代价的”这个想法并不符合直觉。作为人性，它不自然；作为经营业务，它也当然不自然。

[01:01:41] [Lenny]

English:

I love that example of the real estate platform where it's very viscerally, clearly cost. They lost because they didn't roll out experiments to this group for a long time. Such a good example of this idea in action. You mentioned star ratings. I know you spent a lot of time on designing rating systems. Sorry, I didn't mean to imply star ratings. That's just one implementation. Rating systems in general. So maybe just to keep it focused, say a marketplace founder is trying to decide and design how they do ratings, and reviews, and things like that. What's a couple pieces of advice you'd give them for how to do this correctly? And is there a model marketplace you'd point them to like, "These guys really do it really well"? And I know it's super specific based on the marketplace, but is there one just like, "They really nailed it"?

中文翻译:

我很喜欢那个房地产平台的例子，它非常直观、清晰地展示了成本。他们因为长期没有向这组人推广实验而蒙受了损失。这是该理念付诸实践的一个极好例子。你提到了星级评价。我知道你花了很多时间设计评价系统。抱歉，我不是指星级评价，那只是一种实现方式，我是指广义的评价系统。为了保持焦点，假设一位市场创始人正在决定和设计他们如何做评分、评论之类的事情。你会给他们哪几条建议来正确地完成这件事？有没有一个模范市场你可以指给他们看，比如“这家真的做得很好”？我知道这取决于具体的市场类型，但有没有哪一家是“真的做绝了”的？

[01:02:30] [Ramesh Johari]

English:

Oh man, that's a tough one. I think I'll answer the second part first. I don't feel like anyone's really nailed this. Yeah, I think there's a lot of innovation that's happened, but I think fundamentally, we're still playing with the same kind of tools that we had when eBay and Amazon first started thinking about how to do rating systems ages ago. And part of the reason we haven't nailed it is because there's a lot of dynamics in play that lead to what's called rating inflation, where if you look at ratings over time in the marketplace... One of my colleagues, John Horton, who was a professor at MIT and has worked very closely with Upwork, we worked together when I was at oDesk, he was the staff economist there. He's written a couple of really nice papers with this empirical phenomenon that over time, you see the median rating inflating, let's say on marketplaces like oDesk, like Uber, like any of these.

And there's a lot of reasons for this, but one of them is just that there's a reciprocity issue, which is it's effectively, from your perspective, it's kind of costless if someone says to you, "Hey, please leave me a nice rating." And if you're seeing them or you're interacting with them, most people don't want to be mean. So that happens. But there's another aspect of it, which is norming. As the ratings in the marketplace go up, they get normed, so that now you're in a condition, you're like, "A four star rating. I'm really screwing this person over." Whereas maybe when the marketplace started, you didn't think that.

So definitely one thing that we worked on in our research was to think about renorming, the meaning of some of these labels. And renorming could mean something like rather than the star ratings just being poor to excellent, the top rating has actually exceeded expectations. You could go one step further and you could say, "How did this compare to this experience you had in the past that you rated really highly?" And Airbnb had something like this in place, where they would actually ask you to compare, or ask you questions about expectations. I find that that's really valuable because it's easier for people to say, "That was good but didn't exceed my expectations. That was good, but definitely not better than this amazing stay I had two months ago," than it is to say, "Well, I'm going to ding this person and give them four stars." So that's one issue.

And I think another thing I want to point out for any marketplace founder is that something you want to be really careful about is the concept of averaging and whether are the implications of averaging. And that's because a default for many marketplaces is to just average the ratings that people get. It feels very natural, right? Lenny's got five ratings, let me average them. And that actually has some pretty important distributional consequences for the marketplace. Distributional in the sense of who wins, who loses. And that's because if you're averaging and you're really established on a platform, think of a restaurant on Yelp with 10,000 reviews, it's irrelevant what the next review is. It doesn't matter. Nothing's moving it at that point. If you're new and you break into that market, and your first review is negative, you might be completely screwed. In fact, there's some early work on eBay that showed that if your first rating's negative, that could actually immediately cause an 8% hit on your immediate expected revenue, say nothing of long-term consequences. Subsequent work has found that that's a significant indicator of potential exit from the platform, just because now it's very hard to find work. And some platforms do things like maybe they won't show your ratings until you've accumulated a few.

But in the end, this kind of distributional fairness aspect of averaging is pretty significant. And one of the recent papers that we've written is trying to get platforms to think a little bit about that. There's ways to address that interestingly, through the same concept of a prior. And the prior basically says hey, if someone comes into the marketplace and instead of averaging them, I average them together with a prior belief, then maybe what that prior belief does, it says, "Yeah, you got one negative rating, but maybe you got a little bit unlucky," and maybe my prior belief is something which actually pulls your rating up a little bit and allows me to still have you alongside others in the marketplace to give you a chance at getting work, getting rides, etc. So I believe pretty strongly in this kind of distributional fairness element of designing rating systems. I think it's been understudied. And I'll say in general actually, I think rating systems are understudied, which to me is astonishing. Because the biggest change from those Agoras and Trajan's Market elements of those kinds of markets, to me the biggest change is that we get to see what happened with our matches. So as a data scientist working on marketplaces, I feel like it's incredible that more of us don't spend our time thinking about what we're learning from the matches, and what these rating systems are telling us, and what the impact of that is on who wins and who loses in these markets, kind of thinking about the social implications of these things. So that's something I'm pretty passionate about.

中文翻译:

噢，伙计，这很难回答。我想先回答第二部分。我不觉得有谁真正“做绝了”。是的，虽然有很多创新，但我认为从根本上说，我们仍然在使用 eBay 和亚马逊很久以前开始思考评价系统时所用的那一套工具。我们没能做好的部分原因在于，有很多动态因素会导致所谓的“评价通胀”（Rating inflation）。如果你观察市场中随时间变化的评价……我的同事 John Horton（MIT 教授，曾与 Upwork 密切合作，我在 oDesk 时他是那里的首席经济学家）写了几篇非常棒的论文，描述了这种经验现象：随着时间的推移，你会看到中位评分在不断上升，无论是在 oDesk、Uber 还是其他任何平台。

这有很多原因，其中之一就是“互惠问题”（Reciprocity issue）。从你的角度来看，如果有人对你说“嘿，请给我个好评”，这基本上是无成本的。如果你见到了他们或与他们有互动，大多数人都不想表现得很刻薄。所以这种情况就会发生。但另一个方面是“规范化”（Norming）。随着市场评分的上升，它们会被规范化，以至于你现在处于这样一种状态：你觉得“给四星评价就是在坑这个人”。而也许在市场刚开始时，你并不这么想。

所以我们在研究中确实致力于思考“重新规范化”这些标签的含义。重新规范化可能意味着，不再仅仅是“从差到优”的星级评分，最高评分实际上是“超出预期”。你可以更进一步，问：“这与你过去评价很高的一次体验相比如何？”Airbnb 就有类似的做法，他们实际上会要求你进行比较，或者询问关于预期的问题。我发现这非常有价值，因为对人们来说，说“这很好但没超出我的预期”或者“这很好，但绝对没有我两个月前那次超棒的入住好”，要比说“我要给这个人差评，只给四星”容易得多。这是第一个问题。

我想为任何市场创始人指出的另一件事是，你需要非常小心“平均值”(Averaging)的概念及其影响。这是因为许多市场的默认做法就是对人们获得的评分取平均值。这感觉非常自然，对吧？Lenny 有五个评分，让我取个平均值。但这实际上对市场产生了一些非常重要的“分配后果”(Distributional consequences)，即谁赢谁输。这是因为如果你在取平均值，并且你在平台上已经非常资深了（想象一下 Yelp 上一家有 10,000 条评论的餐厅），下一条评论是什么根本无关紧要，它无法撼动任何东西。但如果你是新人，刚进入市场，你的第一条评论是负面的，你可能就彻底完蛋了。事实上，eBay 的早期研究表明，如果你的第一个评分是负面的，实际上会立即导致你预期收入下降 8%，更不用说长期后果了。随后的研究发现，这是潜在退出平台的显著指标，因为现在很难找到活干。一些平台会采取措施，比如在你积累一定数量的评分之前不显示评分。

但归根结底，平均值的这种“分配公平性”方面是非常重要的。我们最近写的一篇论文就是试图让平台思考这一点。有趣的是，可以通过同样的“先验”概念来解决这个问题。先验基本上是说：嘿，如果有人进入市场，我不是直接取平均值，而是将他们与一个先验信念一起取平均值。那么这个先验信念的作用可能是说：“是的，你得到了一个负面评价，但也许你只是运气有点不好。”我的先验信念实际上会把你的评分拉高一点，让我仍能让你与市场中的其他人并列，给你一个获得工作、获得行程的机会。所以我非常坚信设计评价系统中的这种分配公平性元素。我认为这方面的研究还不够。总的来说，我认为评价系统的研究都不够，这对我来说很惊讶。因为从古希腊集市和图拉真市场到现代市场，最大的变化在于我们能看到匹配之后发生了什么。作为一名研究市场的数据科学家，我觉得不可思议的是，我们中竟然没有更多的人花时间思考我们从匹配中学到了什么，这些评价系统告诉了我们什么，以及这对市场中谁赢谁输的影响——也就是思考这些事情的社会影响。这是我非常热衷的事情。

[01:07:14] [Lenny]

English:

I also led the review system flows for a while at Airbnb, and one of the things I'm most proud of is launching what we call double-blind reviews where you don't see the other person's review until you leave your review. The intention was to create more honesty and more accurate reviews. It turned out the biggest impact was review rate went up, because people get this email, "Ramesh left you a review. If you want to see it, should leave a review." And that really increased review rate, which gave us more data. And it was a really fun experiment to work on.

中文翻译:

我也曾在 Airbnb 领导过一段时间的评价系统流程，我最自豪的事情之一就是推出了我们所谓的“双盲评价”(Double-blind reviews)，即在你留下评价之前，你看不到对方的评价。初衷是创造更诚实、更准确的评价。结果发现最大的影响是评价率上升了，因为人们会收到邮件说：“Ramesh 给你留了评价。如果你想看，就得先留个评价。”这极大地提高了评价率，从而给了我们更多的数据。那是一个非常有趣的实验。

[01:07:44] [Ramesh Johari]

English:

There's a great concept in the literature on rating systems called the sound of silence, which is the idea that there's a lot of information in ratings that are not left. So Steve Tadelis, who's a professor at Berkeley, he had a really nice paper with some folks at eBay talking about what they called effective percent positive, where rather than normalizing just by the ratings, they normalized by including ratings that weren't left. And what you found was this was much more predictive of downstream performance of a seller. So there's a lot of information in that lack of a response. So it's cool that you're able to get more of that out.

中文翻译:

评价系统文献中有一个很棒的概念叫“沉默之声”(The sound of silence)，即那些没有留下的评价中包含大量信息。伯克利大学的教授 Steve Tadelis 与 eBay 的一些人写了一篇非常好的论文，讨论了他们所谓的“有效好评率”(Effective percent positive)。他们不再仅仅根据已有的评分进行归一化，而是将未留下的评价也纳入归一化。结果发现，这比单纯的好评率更能预测卖家的后续表现。所以，缺乏回应中包含着大量信息。你能挖掘出更多这样的信息真的很酷。

[01:08:23] [Lenny]

English:

So much easier just to not leave a review than leave a bad review. Right? The downside to you is just much better. Oh man, marketplaces are so fascinating. I could see why a founder would want to be a marketplace founder, because it's just such an interesting space. And hearing your feedback of, no, you're not a marketplace founder. Let's think about the problem you're solving. And it might be a marketplace, might change people's minds. Also, I feel like there's a podcast episode in every topic we touched on. I know we just scratched the surface a lot of things. I know you got to run. Before we get to our lightning round, is there anything else you wanted to highlight, touch on, leave people with that are maybe working on marketplaces, thinking about a marketplace?

中文翻译:

不留评价确实比留差评容易得多，对吧？对你来说，不留评价的负面影响要小得多。噢伙计，双边市场太迷人了。我能理解为什么创始人想成为市场创始人，因为这确实是一个非常有趣的领域。而听到你的反馈——不，你不是市场创始人，让我们思考你正在解决的问题，那可能是一个市场——这可能会改变人们的想法。此外，我觉得我们触及的每个话题都能单独做一集播客。我知道我们只是蜻蜓点水。我知道你得赶时间。在进入闪电轮 (Lightning round) 之前，你还有什么想强调、提及或留给那些正在做市场或考虑做市场的人的话吗？

[01:09:01] [Ramesh Johari]

English:

I think one of the high level points I would make, and like you said, there's an entire podcast in this topic, is that I think people want to imagine LLMs and AI driven data science automating out large parts of what it means to do data science in industry. And I think that's probably the wrong perspective. In some mundane sense, that's true. It's easier for me to code than it used to be before. It's easier for me to develop visualizations than it used to be. I can make dashboards faster. So programmatically, I think it's true in some basic sense.

But what I believe pretty strongly, and I teach data science here, and my students are asked to use LLMs and generative AI on a weekly basis on all their assignments. So I've got an up close and personal beat on this, but I believe very strongly actually is what AI has done for us is it's massively expanded the frontier of things we could think about our problem, hypotheses we could have, maybe things we could test. It's just an astronomical explosion of explanations, and ideas, and principle. And I really think actually what that does is puts more pressure on the human, not less. I think it becomes more important for humans to be in the loop in interacting with these tools to drive the funneling down process of identifying what matters, at all levels. That ranges from you're carrying out a data scientific analysis, and now because you've got these tools, you can hypothesize 10 explanations, maybe 100 explanations. Which of those are you going to focus attention on? What are you going to tell other people to focus their attention on? To you're running experiments, used to have 10 creatives you're testing for a marketing campaign, you got 1,000 creatives, you're testing for that marketing campaign. Maybe that completely changes the game of what it means to run an experiment. What are you actually looking for now? How do you evaluate that

you found something that was good enough? And I think these questions are not getting enough attention. I think people are looking for the automated tool that really cuts the human out. But what I've seen so far, and again, who knows? By 2024, I might have a totally different answer for you. I don't think so. But at the moment, what I see is that humans have actually become far more important to the productive data science loop, not far less.

中文翻译:

我想提的一个高层次观点是（正如你所说，这个话题可以单独做一集播客）：我认为人们倾向于想象大语言模型（LLMs）和AI驱动的数据科学会自动化工业界数据科学的大部分工作。我认为这可能是错误的观点。在某些平凡的层面上，这是真的。我现在写代码比以前容易了，开发可视化比以前容易了，做仪表板也更快了。所以从程序化角度看，这在某种基本意义上是真的。

但我非常坚信的一点是（我在斯坦福教数据科学，我的学生每周的所有作业都被要求使用LLM和生成式AI，所以我对此有非常近距离的观察）：AI为我们所做的，实际上是极大地扩展了我们思考问题、提出假设、进行测试的边界。它带来了解释、想法和原则的天文数字般的爆炸。我真的认为，这实际上给人类带来了更大的压力，而不是更小。我认为，人类在与这些工具互动中保持“在回路”变得更加重要，以驱动“漏斗式筛选”过程，识别出在各个层面上真正重要的东西。这包括：你正在进行数据科学分析，现在因为有了这些工具，你可以假设10种解释，甚至100种解释。你会把注意力集中在哪些解释上？你会告诉别人把注意力集中在哪些解释上？再比如你正在运行实验，以前你为营销活动测试10个创意，现在你可以测试1000个创意。这可能彻底改变了运行实验的游戏规则。你现在到底在寻找什么？你如何评估你找到了足够好的东西？我认为这些问题还没有得到足够的重视。人们在寻找能真正把人踢出去的自动化工具。但到目前为止我所看到的（谁知道呢，到2024年我可能会给你一个完全不同的答案，虽然我不这么认为），我看到的是人类在高效的数据科学循环中实际上变得更加重要了，而不是更不重要。

[01:11:16] [Lenny]

English:

Such an important point. I feel like we need to add AI corner to this podcast where we always think about, how does AI impact what we're talking about on this podcast?

中文翻译:

非常重要的一点。我觉得我们需要在这个播客里加一个“AI角落”，专门思考AI如何影响我们讨论的话题。

[01:11:23] [Ramesh Johari]

English:

Yeah, I can see that. I totally see that.

中文翻译:

是的，我能理解。完全赞同。

[01:11:25] [Lenny]

English:

Okay, we might start doing that. Ramesh, with that, we've reached a very exciting lightning round. I've got six questions for you. Let's try to knock through them so you can go teach your class. Are you ready?

中文翻译:

好，我们可能真的会开始这么做。Ramesh，说到这，我们进入了非常令人兴奋的闪电轮。我有六个问题问你。让我们试着快速过一遍，好让你去给学生上课。准备好了吗？

[01:11:35] [Ramesh Johari]

English:

I am ready.

中文翻译:

准备好了。

[01:11:36] [Lenny]

English:

All right. What are two or three books you've recommended most to other people?

中文翻译:

好的。你向别人推荐最多的两三本书是什么？

[01:11:40] [Ramesh Johari]

English:

When it comes to books, I have one I love that I start with always, which is *How to Lie with Statistics*. It's a tiny book, Darrell Huff from 1954, which is just for anyone that likes data at any level, it's such a fun read. It's a great book. The second thing I recommend to people, and actually this is true even for people who are not expert, is David Freedman was a statistician at Berkeley who passed away in the 2000s, early 2000s. His writing was fantastic in getting us to think hard about process. He was especially fond of what he called *shoe leather statistics*, where you rolled your sleeves up, you got on the ground, boots on the ground, really getting in there, really trying to understand your data.

His writing is fantastic, his explanations are fantastic. He has a few different books at different levels I think people would love reading. Most importantly, what I like about it is he puts such emphasis on driving evidence and understanding of your processes that generate data. And I find often, data scientists don't even look at examples. So at oDesk, it meant are you looking at actual jobs, and what's actually going on in your product before you're trying to do data science on it? So I think that's a Freedman insight, Freedman mantra, and so his writing is great. The last one I was going to mention has nothing to do with data science or anything. It's called *Four Thousand Weeks* by Oliver Burkeman. I'm not a huge self-help type person, but I really like this book a lot. I think it's a little bit stoic in its approach, like stoic philosophy. But the basic point is you're only on earth somewhere in the neighborhood of 4,000 weeks, and my wife and I have this term we call *infinite Q*, which is no matter what you think you get done on a given day, more stuff's going to just keep coming in.

And he basically says that recognizing that is liberating. Because once you recognize it, it doesn't matter what you do. You're always going to have too much to do. There's no point in stressing out about having too much to do. And just that small shift of mindset than puts a lot more attention on the usual thing people worry about, which is, where do I want to prioritize my time? So he has a great way of writing about it, some concrete rules of thumb to help manage that way of thinking. And yeah, I think it's a great book.

中文翻译:

说到书，我有一本总是作为开场推荐的最爱，那就是《统计数字会撒谎》(How to Lie with Statistics)。这是达莱尔·哈夫 (Darrell Huff) 在 1954 年写的一本小书，对于任何层面的数据爱好者来说，这都是一本非常有趣的读物。这是一本伟大的书。

第二本我推荐给人们的（实际上即使是专家也适用）是戴维·弗里德曼 (David Freedman) 的作品。他是伯克利大学的统计学家，于 2000 年代初去世。他的写作在引导我们深入思考过程方面非常出色。他特别喜欢他所谓的“走破鞋底的统计学”(Shoe leather statistics)，即卷起袖子，实地考察，真正深入其中去理解你的数据。

他的文笔极佳，解释也非常精彩。他有几本针对不同水平读者的书，我觉得人们会喜欢读。最重要的是，我喜欢他强调寻找证据和理解产生数据的过程。我发现数据科学家经常甚至不看具体的例子。在 oDesk，这意味着在尝试对其进行数据科学分析之前，你是否查看了实际的职位，以及你的产品中实际发生了什么？我认为这是弗里德曼的洞察和准则，他的书很棒。

最后一本我要提到的与数据科学无关，叫《四千周》(Four Thousand Weeks)，作者是奥利弗·伯克曼 (Oliver Burkeman)。我不是那种特别迷恋成功学的人，但我非常喜欢这本书。我觉得它的方法有点像斯多葛学派 (Stoic philosophy)。核心观点是：你在地球上大约只有 4000 周的时间。我和妻子有一个词叫“无限队列”(Infinite Q)，即无论你认为自己某天完成了多少工作，总会有更多的事情源源不断地进来。

他基本上是说，意识到这一点是一种解脱。因为一旦你意识到这一点，你做什么就不再那么纠结了。你永远会有做不完的事，为事情太多而感到压力是没有意义的。这种心态的小小转变会让你把注意力集中在人们通常担心的事情上，即：我想把时间优先花在哪里？他写得很好，提供了一些具体的经验法则来帮助管理这种思维方式。是的，我认为这是一本很棒的书。

[01:13:52] [Lenny]

English:

What is a favorite recent movie or TV show?

中文翻译:

最近最喜欢的电影或电视剧是什么？

[01:13:55] [Ramesh Johari]

English:

I am a climber, and one movie that I really liked was The Alpinist. I know a lot of people have seen Free Solo, but for anyone that kind of likes that genre, I would recommend they watch The Alpinist. I think climbing is an interesting sport because has very much a psychological aspect of it. And I think that movie is pretty good at this meta level where you reflect a little bit on, what does it mean to make a movie about people who are obviously putting themselves into such risky situations? So I really enjoyed that. On TV, we've been watching Only Murders in the Building, but I'm enough episodes behind right now that I probably won't say anything more, because I am trying to avoid any spoilers and I'm sure there's people out there trying to do the same. So great show though on Hulu.

中文翻译:

我是一名攀岩爱好者，我非常喜欢的一部电影是《阿尔卑斯人》(The Alpinist)。我知道很多人看过《徒手攀岩》(Free Solo)，但对于喜欢这类题材的人，我推荐看《阿尔卑斯人》。我认为攀登是一项有趣的运动，因为它具有很强的心理层面。我觉得这部电影在元层面 (Meta level) 做得很好，让你反思：为那些显然将自己置于如

此危险境地的人拍摄电影意味着什么？所以我很喜欢。电视方面，我们一直在看《大楼里只有谋杀案》(Only Murders in the Building)，但我现在落后了好几集，所以可能不会多说，因为我在努力避免被剧透，我相信也有其他人也在做同样的事。不过这是 Hulu 上的一部很棒的剧。

[01:14:39] [Lenny]

English:

What's a favorite interview question that you like to ask candidates that you're hiring?

中文翻译:

在招聘时，你最喜欢问应聘者什么问题？

[01:14:43] [Ramesh Johari]

English:

I interview people probably that are a little bit different than most of your podcast listeners. But that said, there's one question I like to ask a lot, and that's if you imagine... Often in our interviews in academia, whether it's grad students or faculty will ask people about their plans. And what I like to ask people is, "Okay, now imagine everything works out, all the challenges you're facing work out, all your plans work out, everything hits the top end of your vision for what this could be. What do you imagine is the impact of having done that? Who's being impacted by that? Why is that a big deal that happened?"

And I find that's a really valuable question to ask, because first of all, many people haven't thought about that. We're so short-term focused, we don't even think, "Boy, if everything worked out, what would be the big deal because of what I did?" Startup founders tend to be better at this than most people obviously. But another reason I like it is because you will find in that conversation that their vision expands a little bit of additional spheres that are touched or impacted by what they're thinking about doing. So on both sides, it's kind of a revealing question, I think. So I find that important for my line of work, but my hunch is that might be useful for some of your listeners too.

中文翻译:

我面试的人可能与你播客的大多数听众有点不同。尽管如此，有一个问题我非常喜欢问，那就是让你想象……在学术界的面试中，无论是研究生还是教职员，我们都会询问他们的计划。我喜欢问的是：“好，现在想象一切顺利，你面临的所有挑战都解决了，你所有的计划都实现了，一切都达到了你对这件事愿景的最顶端。你想象这样做会产生什么影响？谁会因此受到影响？为什么这件事的发生是一件了不起的大事？”

我发现这是一个非常有价值的问题。首先，很多人还没想过这个问题。我们太专注于短期了，甚至不会去想：“天哪，如果一切顺利，我所做的事情会有什么了不起的意义？”初创公司创始人显然比大多数人更擅长这一点。但我喜欢这个问题的另一个原因是，你会发现在对话中，他们的愿景会扩展到一些额外的领域，这些领域会被他们正在考虑做的事情所触及或影响。所以我认为，从两方面来看，这都是一个很有启发性的问题。我觉得这对我这一行很重要，但我直觉它对你的一些听众也可能有用。

[01:16:15] [Lenny]

English:

Yeah, such a unique perspective on interviewing, versus most of the guests that I interview in tech company.

中文翻译:

是的，这确实是一个非常独特的面试视角，与我采访的大多数科技公司嘉宾不同。

[01:16:03] [Ramesh Johari]

English:

Yeah, normally there's a coding question, right? I should say I would never ask a coding question post November 2022 after we got AI to help us code. I think it's a superpower.

中文翻译:

是啊，通常会有一个编程题，对吧？我得说，在2022年11月我们有了AI辅助编程之后，我再也不会问编程题了。我认为AI是一种超能力。

[01:16:15] [Lenny]

English:

AI corner. What is a favorite product you've recently discovered that you really like?

中文翻译:

又是“AI角落”。你最近发现并非常喜欢的某个产品是什么？

[01:16:22] [Ramesh Johari]

English:

I also really like cycling. And I'm not ashamed to admit that I think that e-bikes are the greatest thing for cycling. Admittedly, I'm late forties, so maybe I'm the right target demographic too. But yeah, I love my e-road bike. It's great, because it's not one of those with a throttle, you have to work, but it kicks in just when you're on your sixth hill and you don't want to go up the last hill anymore on the way home. So yeah, that's amazing. I think that's just transformative for people that like cycling, but have busy lives. And I think another one that my son who's 10 roped me into actually, is we were in Santa Cruz browsing at a kitchenware shop of all places, and he saw an outdoor pizza oven, a tiny portable one. And he just did research for two weeks and insisted we get one. So he got one over the summer, and after we got it, he refused to eat pizza out anymore as a 10-year-old. So maybe that's the best thing I could say about the quality of pizza you can get from a home outdoor portable pizza oven.

中文翻译:

我也非常喜欢骑行。我不羞于承认，我认为电助力自行车（E-bikes）是骑行界最伟大的发明。诚然，我快五十岁了，所以可能也是正确的目标受众。但我真的很爱我的电助力公路车。它很棒，因为它不是那种带油门的，你还是得蹬车，但当你爬到第六个山坡，在回家路上再也不想爬最后一个坡时，它就会介入。所以，这太神奇了。我认为对于喜欢骑行但生活忙碌的人来说，这是革命性的。

另一个是我10岁的儿子拉我入坑的。我们在圣克鲁斯逛一家厨具店时，他看到了一个室外披萨烤炉，是一个很小的便携式的。他做了两周的研究，坚持要买一个。于是他在夏天得到了一个。买回来之后，作为一个10岁的孩子，他再也不愿意去外面吃披萨了。这大概是对家用室外便携披萨烤炉能做出的披萨质量最好的评价了。

[01:17:20] [Lenny]

English:

Oh my God, I'm hungry. I am going to go have to get some pizza now. What is a favorite life motto that you like to repeat to yourself, share with folks, find useful in your day-to-day?

中文翻译:

噢天哪，我饿了。我现在得去弄点披萨吃。你最喜欢的人生格言是什么？那种你会对自己重复、与人分享、并在日常生活中觉得很有用的格言。

[01:17:31] [Ramesh Johari]

English:

A lot of my work involves talking to students of all stripes. And I guess these students go on to be data scientists, go on to be founders, and a lot of them go in the tech industry. So maybe in that sense, that advice is relevant. My main thing I tell people is slow down. I think what I've found has been happening, is we're so convinced that speed is the way you're going to find the right answer, that I just don't think we slow down to develop meaningful mental models of the things we're doing. That's certainly true in the research projects I work on. It's consistently true when I talk to people in business, and I ask them about their... By mental model, I just mean if you're running a marketplace, what is your model of what people care about? What makes people stay versus leave? What makes matches work versus not work? All those things shape a roadmap in your mind. And I think a lot of roadmapping, a lot of execution, paper writing in academia has all just become far more fast-paced, at the expense of deeper thinking about these kinds of structural features of the thing you're building. So with my students, but also I think with people I interact with in industry, I think slowing down is actually more of a virtue than it's given credit for.

中文翻译:

我的很多工作涉及与各色学生交流。这些学生以后会成为数据科学家、创始人，很多人会进入科技行业。所以从这个意义上说，这条建议是相关的。我告诉人们最主要的一点是：慢下来。我发现现在的情况是，我们太坚信速度是找到正确答案的唯一途径，以至于我觉得我们没有慢下来为我们正在做的事情建立有意义的“心理模型”（Mental models）。在我从事的研究项目中确实如此。当我与商界人士交谈，询问他们的……我所说的心理模型是指：如果你经营一个市场，你对人们关心什么的认知模型是什么？是什么让人们留下而不是离开？是什么让匹配成功而不是失败？所有这些东西构成了你脑海中的路线图。我认为现在的路线图制定、执行，甚至学术界的论文写作，都变得节奏太快了，代价是缺乏对你所构建事物的结构性特征的深入思考。所以，无论是对我的学生，还是对我在工业界接触的人，我认为“慢下来”实际上是一种被低估的美德。

[01:18:50] [Lenny]

English:

Very similar to a motto that a recent guest shared, which I think was go slow to go fast, or stay smooth to go fast.

中文翻译:

这与最近一位嘉宾分享的格言非常相似，我想那是“以慢为快”（Go slow to go fast），或者“保持平稳才能快”。

[01:18:59] [Ramesh Johari]

English:

Yeah, I like that. Maybe I'll pilfer that, when I go talk to my grad students [inaudible 01:19:03].

中文翻译:

是的，我喜欢那个。也许我去跟我的研究生谈话时会借用一下。

[01:19:04] [Lenny]

English:

Final question. You're a professor at Stanford University, which sounds incredibly cool. What's something about being a professor at Stanford in particular or in general that would surprise people, either good or bad?

中文翻译:

最后一个问题。你是斯坦福大学的教授，这听起来酷毙了。关于在斯坦福当教授，或者广义上的教授生活，有什么会让人们感到惊讶的事情吗？无论是好的还是坏的。

[01:19:17] [Ramesh Johari]

English:

Yeah. I mean, we've had a rough ride, as everybody probably knows. Stanford's been in the news for a lot of not so great reasons, I think over the last five years especially. So I don't know if this is the right kind of surprise, but I think one thing that I find really energizing at Stanford is people have never asked me for credentialing here. And what I mean by that is that I came from a bunch of other good schools, and obviously I've spent time in industry with a lot of great companies. And a kind of cultural dynamic that can often develop is, "Well, before I'm going to talk to you, I want to know something about why you're worth talking to. Give me your credentials. Where are you a grad student or where are you a professor? Tell me about yourself first."

One of the things that I found very surprising when I came here is just how that never happened at any level. Grad students tell me this all the time. Go talk to someone across campus and just launch right into a conversation about how your X meets my Y, and we have something we could do together. As a faculty member, it happens all the time. I just had a conversation a couple days ago with someone about effectively a marketplace of experiment designs for nano fabrication here, which is totally out of left field for things I do, and yet seamless. Our conversation was about the substance rather than the credentialing. I really think part of the reason for that is that Stanford is sort of unique in that it doesn't have a weakness across the board. We have strong professional schools, law, business, medicine, strong engineering schools, strong humanities and social sciences. And then that and the weather is what I usually tell people honestly, which matters a lot. People are willing to walk anywhere. I think those things combine to create a culture and an environment where you don't credential everybody. And I think that means a lot. I think that's something that I haven't found elsewhere. And if people wanted to know something about what's Stanford's like on the inside, I think that's one aspect of it that probably isn't discussed very much. I think that's part of what makes it really fun to be here.

中文翻译:

是的。我的意思是，正如大家可能知道的，我们经历了一段艰难的时期。斯坦福在新闻中出现了很多不太好的原因，尤其是在过去的五年里。所以我不知道这是否是那种“惊喜”，但我认为在斯坦福让我感到非常有活力的一点是：这里的人从未要求我提供“资历证明”（Credentialing）。我的意思是，我来自其他几所名校，显然我也在工业界与很多伟大的公司共事过。通常会形成的一种文化动态是：“好吧，在我跟你交谈之前，我想知道你为什么值得我交谈。给我看看你的资历。你在哪儿读的研究生？你在哪儿当教授？先介绍一下你自己。”

当我来到这里时，我发现非常惊讶的一点是，这种情况在任何层面上都没有发生过。研究生们经常告诉我：去校园另一头找个人聊天，直接就开始讨论你的研究如何与我的研究结合，以及我们可以一起做些什么。作为教职员，这种情况也一直在发生。就在几天前，我还和某人讨论了一个关于纳米制造实验设计的“市场”，这完全超出了我的研究领域，但交流却非常顺畅。我们的谈话是关于实质内容的，而不是关于资历的。我真的认为部分原因是斯坦福在各个领域几乎没有弱项。我们有强大的专业学院（法律、商学、医学），强大的工程学院，强大的人文和社会科学。除此之外，老实说，还有天气，这很重要，人们愿意到处走动。我认为这些因素结合在一起，创造了一种不看人下菜碟的文化和环境。我认为这意义重大，这是我在其他地方没发现的。如果人们想了解斯坦福内部是什么样子的，我认为这是可能不常被讨论的一个方面。这也是在这里工作非常有趣的原因之一。

[01:21:27] [Lenny]

English:

It's also an incredibly dreamy campus, that is very joyful to walk around. That helps, I'm sure. Ramesh, I feel like we got people's brains tingling. I think we've created new marketplace founders, and also convinced people maybe they aren't marketplace founders. So maybe we netted out zero new marketplace founders. Two final questions. Where can folks find you online if they want to reach out? And how can listeners be useful to you?

中文翻译:

那也是一个梦幻般的校园，走在里面非常愉悦。我确信这也有帮助。Ramesh，我觉得我们让听众的大脑兴奋起来了。我想我们创造了一些新的市场创始人，同时也说服了一些人也许他们并不是市场创始人。所以也许我们新增的市场创始人是零。最后两个问题：如果人们想联系你，可以在哪里找到你？听众能为你做些什么？

[01:21:49] [Ramesh Johari]

English:

I think the easiest way, if someone's interested more on the industrial side is probably LinkedIn. You send me a message or connect there. Also, because I'm an academic, I have my own Stanford webpage, and it's pretty easy to figure out how to find me there as well. And how can listeners help me? I kind of feel the most important thing that someone listening to this could do is take forward some of the messages that came out in terms of what it means to be data literate. And I think there's a lot you can do to educate yourself there.

Maybe one final thought I'll share is that in the same way that AI generates a lot of ideas, AI also generates a lot of prose. And in data science, that can actually be deadly because you're getting more explanations that sometimes maybe are extraneous. So taking that as a little vignette, I think that what the world needs is data literacy on the part of people interacting with these tools and with each other. So that's the thing I care most about. The things I teach, the things I do research on, they're all connected to that theme. And so that's where I'm pretty excited. I do work with companies regularly, and so if there's interesting opportunities that fall in the sphere of stuff we've discussed on the podcast, always happy to listen.

中文翻译:

我认为最简单的方法，如果有人对工业界方面更感兴趣，可能是 LinkedIn。你可以在那里给我发消息或建立联系。另外，因为我是学术界人士，我有自己的斯坦福网页，在那里找到我也很容易。

听众能如何帮我？我觉得听众能做的最重要的事情，就是把今天传达的一些关于“数据素养”（Data literate）的信息发扬光大。我认为在自我教育方面有很多可以做的。

也许我最后想分享的一点是：就像 AI 生成很多想法一样，AI 也生成很多文字。在数据科学中，这实际上可能是致命的，因为你会得到更多有时可能是多余的解释。所以作为一个小结，我认为世界需要的是人们在与这些工具互动以及彼此互动时的“数据素养”。这是我最关心的事。我教的课、我做的研究，都与这个主题有关。所以我对此非常兴奋。我经常与公司合作，所以如果在我们播客讨论的领域有有趣的合作机会，我总是乐于倾听。

[01:23:00] [Lenny]

English:

Awesome. I think we've made a dent in helping people become a little more data literate. Ramesh, thank you so much for being here.

中文翻译:

太棒了。我想我们在帮助人们提高数据素养方面已经迈出了一小步。Ramesh，非常感谢你能来。

[01:23:07] [Ramesh Johari]

English:

All right. Thank you so much, Lenny.

中文翻译:

好的。非常感谢，Lenny。

[01:23:08] [Lenny]

English:

Bye everyone. 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。非常感谢收听。如果你觉得这期节目有价值，可以在 Apple Podcasts、Spotify 或你喜欢的播客应用上订阅本节目。此外，请考虑给我们评分或留下评论，这能真正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。下期节目见。