

SDS PODCAST

EPISODE 811:

SCALING DATA

SCIENCE TEAMS

EFFECTIVELY, WITH

NICK ELPRIN



- Jon Krohn: 00:00:00 This is episode number 811 with Nick Elprin, CEO of Domino Data Lab. Today's episode is brought to you by AWS Cloud Computing Services.
- 00:00:15 Welcome to the Super Data Science podcast, the most listened to podcast in the data science industry. Each week we bring you inspiring people and ideas to help you build a successful career in data science. I'm your host, Jon Krohn. Thanks for joining me today. And now, let's make the complex simple.
- 00:00:46 Welcome back to the Super Data Science podcast, today the extremely intelligent and extremely successful machine learning entrepreneur, Nick Elprin is our guest on the show. Nick is co-founder and CEO of Domino Data Lab, a colossal Bay Area startup that has raised over \$200 million in venture capital from some of the world's most prestigious VC firms. Prior to co-founding Domino Data Lab 11 years ago, he worked as a technologist at Bridgewater, the well-known hedge fund. He holds both a BA and MS in computer science from Harvard University.
- 00:01:17 Today's episode may appeal most to technical folks, but has tons of content that will be of interest to anyone in or interested in commercializing data science or AI. In this episode, Nick details how organizations can leverage enterprise platforms to efficiently scale their data science teams and data science workflows, the exact team size at which integrating such a platform becomes worthwhile, how to ensure AI projects are commercially successful, the tech stack they use at Domino to create such a performant platform, and his top tip for growing your own colossal data science startup. All right, you ready for an exquisite episode? Let's go.
- 00:02:00 Nick, welcome to the Super Data Science podcast. I'm delighted to have you here. Where are you calling in from today?



- Nick Elprin: 00:02:05 I'm in San Francisco, and we've got a somewhat unusual, nice sunny day.
- Jon Krohn: 00:02:09 Nice. Yeah, you and I were chatting just before we started recording, and I'd recently had a trip to San Francisco. And the entire time I was there it was blue and sunny. I wouldn't say it was hot, but a half hour drive in any direction out of San Francisco and it was scorching. It was over 100 degrees.
- Nick Elprin: 00:02:30 Yeah, well it's maybe a consequence of climate change.
- Jon Krohn: 00:02:35 A factor perhaps, and the El Nino or El Nina effects I think as well are playing a factor in the current years. But amazing city, you and I were talking about how it's buzzing again. The AI boom has kicked off a resurgence that the pandemic had kicked a little bit into San Francisco and it's vibrancy, but I was blown away when I was there. First thing I did when I landed, I went to a coffee shop, and every conversation that I could overhear would be people talking about getting some software sales done, or literally pair programming, sitting there in the coffee shop talking to somebody who's on a laptop saying, "Ah, I don't think you need that kind of tensor representation here." And then in Whole Foods, checking out my food, hearing people talking about the product that they're building in the line behind me, and everywhere I went it was all about AI startups. So, it must be really cool to be deep in the middle of that.
- Nick Elprin: 00:03:38 Yeah it's back, and well, it's particularly nice for me to see, I grew up in San Francisco. I'm a rare San Francisco native, although I spent some time on the East Coast before I moved back here. And yeah, there was a lot of talk of San Francisco's demise a few years ago, and it did get a little rough post pandemic, but it's back and there are lines out the door at the good coffee shops, and the train stations are busy in the morning. And like you

said, you walk into a coffee shop and everyone's got some kind of IDE up on their laptop, and if it's not Java, or Python or whatever, it's something more basic. But yeah, I see lots of text editors and not a lot of web browsing.

- Jon Krohn: 00:04:22 Yeah, it's super cool. I am committed to spending more time around there because of how inspiring it was, for sure. Now, we didn't meet on the West Coast, we met on the East Coast. We met at the NYU, the New York University AI Masterclass. So Jepson Taylor, formerly Ben Taylor, who has been probably the guest that we've had most often on this show of anyone. He-
- Nick Elprin: 00:04:45 He has a special jacket?
- Jon Krohn: 00:04:48 Exactly. We gave him a Super Bowl ring for each appearance, so he's really blinged out now walking around. So yeah, so I hosted a panel on AI at this AI masterclass. There were a lot of great guests on that panel. Daniel Hume was another one of them, he was actually just recently in an episode, number 807, a couple of weeks ago on this show. And yeah, now we've got you, really excited to have you here. Let's dig right into what you're doing at Domino Data Labs, where you're co-founder and CEO. So, the Domino Data Lab promotes a unified and open platform for AI. It emphasizes end-to-end models and data science lifecycle management as well as MLOps automation. That's a lot of different things going on, so maybe you could summarize for us, I'm sure as the CEO this is your bread and butter, to tell us what the pain points are for your users and how the Domino Data Lab platform addresses them.
- Nick Elprin: 00:05:48 Yeah. So, I started building Domino about 10 years ago, and over the last decade we've built what I believe is the most comprehensive and the best platform that gives large organizations, particularly enterprises everything they need to do mission-critical AI at scale. And I think of



that as really having three elements or three facets of what a large organizations need to do data science ML at scale. The first is we make it really easy for data scientists to orchestrate, get access to, weave together all the different infrastructure that you need when you're doing AI work. So you need a lot more compute, you need access to data, you need agility to use all the new latest and greatest software tools, packages. The state of the ecosystem is changing so fast that every week there's a new thing you might want use. And in enterprises, making it easy to get access to all those infrastructure components, all those resources can be very friction full, because IT is putting up bottlenecks and gates, and things like that.

00:07:07 And at the same time, if you're a data scientist you're not a DevOps expert, you're not an infrastructure engineering expert, so you may not want to be setting up your spark cluster and debugging parallel jobs, and things like that. So the first thing we do, we provide self-serve access to all the infrastructure you need to be really productive and experiment rapidly, and stay on the cutting edge. And by the way, we do that in a way that also meets the needs of security conscious IT organizations. So, imagine providing infrastructure access with security controls in place and templates, because again, in an enterprise context you don't want to just give all the data scientists free rein to build a Wild West, or whatever they want. So, that's the infrastructure layer of our stack. And we're, as far as I know, the only advanced data science platform that has native capabilities for multi and hybrid cloud orchestration.

00:08:06 So as data scientists, whether you're running an interactive development experience like Jupyter, whether you're running a batch training job or a realtime inference, you're deploying a model for realtime inference, we can push that to any cloud or on-premise. Okay. So



on top of that, next layer of the stack, once you got all the infrastructure stuff you need, we've built a, like you said, an integrated experience that facilitates the model development lifecycle. I think about this as like what the Microsoft Office suite is for workspace productivity, or what the Adobe Creative Suite suite is for designers, it's all the tools you need to go through your workflow, your productivity apps. So interactive development, experiment management, model deployment, model monitoring, and then critically making that a tight closed loop. So if you go to deploy a model, we'll automatically set up the monitoring rules for you.

00:09:09 If we detect that drift has occurred, we will automatically create a development environment for you with all your raw materials, your code, your data, your software package definitions precisely as they were when you first deployed your model. Because we really believe that data science, it's not a straight line, it's a loop. And the most effective organizations are the ones that speed up this iterative development life cycle. So it's not just deploy a model, it's deploy it, monitor it, and then continuously improve. So, we think a lot about how do we streamline the process, the workflow for data scientists as they go through that whole iterative, continuously improving model development life cycle? So, that's layer two.

00:09:58 And then layer three, which it's one that I think is most exciting and probably most valuable, and it's one that unfortunately doesn't get a lot of attention because it's not really sexy, it is creating the system of record for data science work and artifacts. So the models that you deploy, the history of experiments that you've done, the projects you've tried, keeping that all in one place so there's a single source of truth, and so that organizations can find and build on past work instead of reinventing the wheel, and that organizations can have standards and consistent ways of working. And let me just unpack that



for a little bit. I was talking to an SVP of data science at a big media company a couple weeks ago, and he said his is the only team in the whole company that can onboard a new team member in 24 hours.

00:11:07 And that has been critical to them accelerating their productivity, increasing their productivity. He said over the last couple of years they've seen a 6X improvement in the throughput of the productivity of his data science organization. And he attributed that largely to the ability to reuse past work and stop reinventing the wheel all the time. So a new person joins the team, they're ready to go with, "Hey, here's how we do a project, here's the life cycle. Here's how you access your data," we give you access to the platform. Everything you need is at your fingertips ready to go. You don't have to go hunt around and ask people, and find wikis, and have people send you connection strings or whatever. So I know it was a really long answer, but I thought that'd be good to set the stage a little bit.

00:11:58 So, the way we see the world is what enterprises need to scale mission-critical AI, self-serve access to governed flexible infrastructure, great productivity suite to streamline the cycle of working through the model development life cycle for data scientists, and then the system of record that helps organizations build on past work instead of continuously reinventing the wheel.

Jon Krohn: 00:12:22 Sometimes we have people talk for half an hour without giving me a chance to say anything at all, so that wasn't the longest answer ever, and it makes perfect sense. You also did my job for me at the end there by summarizing what you said there. So the idea in general, being the Domino Data Lab is comprehensive in your thoughts, unbelievably as the CEO, the best platform-

Nick Elprin: 00:12:42 Shocking, right? I know.



- Jon Krohn: 00:12:43 ... of mission-critical AI at scale, particularly for large enterprises. And you talked about those three layers, you just summarize them with that infrastructure layer, so I'm guessing that allows us to ... Did you even say it would allow me to do something in something like a Jupyter notebook environment, where it's more like a scripting environment and push that easily into production code? That's what the infrastructure layer allows?
- Nick Elprin: 00:13:03 I think of that more is that middle layer, the model development, life lifecycle, productivity suite. But the way they work together is if you want a Jupyter environment, there's a UI, think of it as like a pallet where you get to pick your different tools. You're not limited to Jupyter. You can say, "Maybe I want Jupyter, maybe I want Jupyter Lab, maybe I want RStudio. So, we've built a tool agnostic platform. We really believe the ecosystem, especially the open source ecosystem, is going to always be the best source of innovation. We don't want to build our own IDE.
- 00:13:38 I think data scientists want flexibility, and I think it's really a mistake for vendors to try to box a data scientist into using a particular tool or framework. But we let you say, "Okay, I want Jupyter, I want Jupyter Lab, I want RStudio." I mean we support, I'll call them legacy analytics applications like MATLAB, or SAS, S-A-S. And then you can say, and this is where the infrastructure piece comes in, you can say, "Well, where do I want to run that? I want to run that on a giant GPU box in AWS," or, "I want to run that on my NVIDIA DGX SuperPOD that I've got on premise." Or, "I don't need GPUs, but I need a ton of memory. Give me half a terabyte of memory," and we'll spin that up for you and link in your Jupyter Notebook, and you get the self-serve provision dev environment right at your fingertips.



00:14:28 You touched on something else, Jon, that maybe I just highlight. We are focused on what we think of as a code first approach to data science. So, the data scientists working in Domino, they are writing code. It could be Python, it could be R, it could be MATLAB, it could be SAS, it could be Scala, but we really believe that the highest value, highest impact applications of AI are not going to be drag-and-drop, point-and-click commoditized. They are going to take the flexibility that you get with coding.

Jon Krohn: 00:15:12 This episode of Super Data Science is brought to you by AWS Trainium, and Inferentia, the ideal accelerators for generative AI. AWS Trainium and Inferentia chips are purpose built by AWS to train and deploy large scale models. Whether you are building with large language models or latent diffusion models, you no longer have to choose between optimizing performance or lowering costs. Learn more about how you can save up to 15% on training costs and up to 40% on inference costs with these high performance accelerators. We have all the links for getting started right away in the show notes. Awesome, now back to our show.

00:15:53 Yep, I see the same thing, we've talked about on the air before with other companies where this idea of the citizen data scientist, does that really exist? Have you ever met one?

Nick Elprin: 00:16:08 Is this a rhetorical question?

Jon Krohn: 00:16:13 Yeah. Well it is, but I guess if you have met one then it would be interesting to hear about it. But there's this idea of this citizen data scientist, which is someone who works in the accounting department, or HR or something, and they use a low code or no code tool to circumnavigate meetings to email or Slack, a data analyst or data scientist, and they can just build a little pipeline

themselves. There's tons of tools out there that do that, and I'm sure if we had someone on from one of these low code or no-code tools they could explain to us exactly where these people are and where they're seeing market traction. But I don't meet a lot of people that fit in that box.

- Nick Elprin: 00:16:57 Look, there's certainly a place for low code or no-code data analysis. I think that those sorts of tools are great for ETL, they're probably great for some BI, so I wouldn't say that all analytics and data work has to be done with code. I think that's not efficient. But I think the mistake that the world or the market got into for a little while, and I think this is driven in part by some of the third party analysts in the space really pushing this idea of citizen data science, was this idea that complex modeling will be done with no-code solutions. I've never seen that work. And I'll tell you about a couple of the use cases that our customers have been doing with our platform, and the Navy is building underwater mine detection models, and they're using lots of different data sources from imagery of the seabed, to sonar data and whatever.
- 00:18:04 And some other stuff I probably don't know about. And I don't know, they're not going to do that with a point and click tool. Or there was one of the big pharmaceutical companies we were working with has developed models so they can, instead of doing a genome sequence to understand certain characteristics of a cell, they built deep learning computer vision models so they can actually look at an image of a cross section of the cell in a slide, like a histopathology image. And by using image recognition instead of a full genome sequence, they can much more rapidly and cost effectively diagnose and track patients. And again, that's not a drag and drop use case. So, I come back to our North Star is enabling mission-critical AI, I think mission-critical AI is going to be done with code, it's not going to be done with drag and

drop. But there are a ton of opportunities to upskill people to help them learn code so they can be effective at building some of these advanced use cases.

00:19:15 Some of the customers we work with, they have role, a persona they've started talking about as a citizen developer. And it's interesting because these folks, they have a statistics background or a hard science background, economics, physics, they're comfortable writing and reading code. They can sling some Python. What they don't know is the full depth of software engineering, and they don't know infrastructure engineering. They couldn't go write a distributed system. But it turns out that if you can abstract away a lot of the infrastructure and DevOps complexity, then you widen your aperture of talent. And you can take these folks who come from statistics hard sciences background know how to read and write some code, and you can make them very effective in contributing to an advanced data science organization.

Jon Krohn: 00:20:15 Sounds a lot like me.

Nick Elprin: 00:20:17 Yeah, what was your background? Sorry, I don't know.

Jon Krohn: 00:20:19 No, so I did a PhD in neuroscience. And in that, taught myself machine learning and statistical computing in order to be able to analyze large datasets in genomics and brain imaging. But I've never taken a computer science course.

Nick Elprin: 00:20:39 Yeah. I don't know how old you are so I'm going to guess, you probably hated if you had to deal with configuring your distributed jobs to work with an MPI framework, or getting the compiler flags right to make sure the BLAST packages in your Unix environment were optimized. So you were dealing with that, that's not what you should be dealing with.



- Jon Krohn: 00:21:01 It's interesting, I think part of what made me so well suited to doing that work as a neuroscientist was that I actually didn't mind that much. So I mean, what attracted me to machine learning and statistical computing as opposed to say, growing tissue cultures in a lab or genetically modifying mice, is that you got typically instant feedback. You knew if something worked, you could examine. And so, futzing around with different arguments, it's cool you know BLAST, that's something I hadn't thought about in a while, but I didn't mind that so much.
- 00:21:48 I think for me it's something not about disposition, just more about that you can't be expert at everything. And so, if I have the opportunity to learn more about a transformer architecture or an LLM, or how I could be doing MLOps, we have lots of MLOps episodes on the show. There's things about it that are interesting, but I'm going to skew towards learning more about models and how they work. Because I don't know, for whatever reason that's the thing that's always fascinating me the most about this is modeling data. And so, the more that things like MLOps, the kind of infrastructure layer that you were describing as your first of three layers at Domino Data Lab, that's something that I'm delighted to have automated for me.
- Nick Elprin: 00:22:34 I mean, I'm glad that you like doing that stuff, but there are a lot of people who don't. So, imagine if we could elevate more of the technical population so that if you did have the skills for analytical reasoning and some basic programming, but you didn't enjoy or want to spend time doing all that DevOps stuff, imagine if we could make you super productive to contribute on some of these advanced teams. That's a lot of what we're doing.
- Jon Krohn: 00:22:59 For sure. It's been many, many years since anyone has taken a line of code I wrote and put it into a production



system. So yeah, I totally get the value of that. I like that idea of the citizen developer. You mentioned a couple of big use cases there for the Domino Data Lab, the US Navy for example. Do you have any other kinds of big logos that you'd like to tell us about, or any particular kinds of use cases maybe that shows the breadth of how big enterprises are using your platform?

- Nick Elprin: 00:23:30 Yeah, so if you come back to this idea of we're focused on enabling mission-critical AI, what that means in practice is really we work with companies where models and AI are typically driving top line of the business, their core to revenue, their core to strategic advantage, or their core to competitive advantage or core to strategy. And so, in practice that means we work largely with financial services organizations, banks, insurance companies, asset managers. We work heavily with the life sciences space. We're working with 10 of the largest 20 pharmaceutical companies in the world right now. And then we're also, as I alluded to, doing quite a bit of work with the federal government, especially in the defense space. And I don't remember quite which customers I can mention by name and which I can't, but I can tell you about some interesting use cases.
- 00:24:28 In pharma and life sciences, I mentioned this image processing use case, but there's also a lot of work happening throughout the full R&D lifecycle. So, early stage we're working with some companies that are doing molecular compound synthesis and molecular compound design. If you can accelerate identification of promising compounds, that can jumpstart the whole process of developing new drugs. And then at the tail end of the process, there's a lot of opportunities for using predictive modeling to optimize clinical operations and shorten the time to complete clinical trials. And that's a really big deal, if you can bring a new drug to market a couple



months faster you're talking about saving lives and making a ton of money for the business.

00:25:32 So, that's interesting. Actually, one customer I know I can talk about, GSK, big pharmaceutical company, they're now using our platform for all of their clinical programming. So, that's the work that data scientists and clinical programmers do to analyze and prepare results of their clinical studies before they go get regulatory approval. And I think that's interesting because that's not GenAI, which I know all anyone ever talks about right now is GenAI, but I think we also got to remember there's a lot of incredibly high value data science that is not generative AI. And I still think that traditional machine learning, and even in a lot of cases traditional statistics is still a lot of the workhorse for some of the highest value models that a lot of companies are using. So, those are some examples in life sciences. In insurance, I mean, obviously modeling is the core of an insurance company's business. How do you assess risk and price policies? That is a modeling question.

00:26:42 And so the better you can do that, the better you can make your prices more competitive, the better you can attract customers, so the better you can manage your own portfolio of risk. And so, we have customers who are using novel ways of assessing risk, and pricing risk, and simulating risk and things like that. But then also with insurance, there are interesting improvements to customer experience. Like if you get into a car accident, it's a really terrible process to go make a claim, have a claims adjuster come out and look at the damage on your vehicle or whatever. And so, there's one use case I liked a lot that's an app where if you're in a car accident, you take a picture of your car and the accident, and models and machine learning will create an estimate of damages and bypass the whole process for having claims adjusters involved.



00:27:46 And I think that's a great example of improving customer experience and increasing operational efficiency for the business. And maybe along those lines of operational efficiency improvements, claims processing or any kind of document processing is another really common use case that we see as a low hanging fruit way to create a ton of value. There's an insurance company we work with that they've now built models that can automatically resolve something like 60% of their claims. And you just think about how many people are involved in that process, and what a terrible experience it can otherwise be for a consumer to have to deal with that. And it's just great. I mean, it's like a radically better win-win for everybody involved. So, those are a few that come to mind.

Jon Krohn: 00:28:43 Do you ever feel isolated, surrounded by people who don't share your enthusiasm for data science and technology? Do you wish to connect with more like-minded individuals? Well, look no further. Super Data Science community is the perfect place to connect, interact, and exchange ideas with over 600 professionals in data science, machine learning and AI. In addition to networking, you can get direct support for your career through the mentoring program where experienced members help beginners navigate. Whether you're looking to learn, collaborate, or advance your career, our community is here to help you succeed. Join Kirill, Hadelin and myself, and hundreds of other members who connect daily. Start your free 14-day trial today at superdatascience.com and become a part of the community.

00:29:27 Yeah, those were great. I'd like to reiterate my support for the statement that there's a lot of AI out there, or a lot of modeling out there that provides value beyond generative AI. It's interesting how in the past few years there have been these mind-blowing improvements in generative AI capabilities that are headline grabbing, but down on the

ground the vast majority of automated applications where we're using data to do some kind of modeling and automate something, that is rarely generative AI based.

00:30:02 There's cool things around, okay, now I can have a natural language interface to something that I didn't have before. It's the main use case. But then you also see people tack that into every product, and I don't need that in every product. I use this thing called WHOOP, which is a heart rate monitor and sleep detector that I wear on my wrist all the time. And the WHOOP added in, I don't know, six months ago, this chat functionality where I can talk in the app and get advice, like exercise advice. And you just never do. It'd be interesting to know if there are power users of that where you're chatting to your heart rate monitor app and getting advice on what you could be doing with exercise. I don't know. I mean, there's probably somebody out there, but it's not a huge value add for probably most users. It's probably a very small percentage that use it a relatively small amount of time.

Nick Elprin: 00:30:55 Totally. Well, and you layer onto that, that we still haven't solved the hallucination problem. And you layer onto that the costs on the backend of LLM inference, and there's a lot of hype and there's a still a lot of nonsense around GenAI, I think. And my hope is that this massive extra amount of attention now that the whole world has on AI will help elevate the more traditional data science ML use cases, but my worry is that it drowns it all out. And honestly, the direction I see more likely right now is the latter, the drowning it out. Because I talk to chief data science officers, chief analytics officers, and what they tell me in confidence, in private is all their CEO or all their board wants to hear about is GenAI, even though that's probably only going to be 20% of what's going to create value in their business. So, yeah.

- Jon Krohn: 00:32:09 Another problem that it's created is because for the natural language in natural language out experience that anyone now can access, including executives at your company, they get this magical feeling experience with that conversation, and then so they make what seems like a very obvious, safe leap that now AI can just do anything automatically where it ... I mean, to build something like GPT-4.0, or Google Gemini, we're talking about hundreds of millions of dollars at a minimum to train something that has that magical capability. And there are clever ways, maybe again, if we're talking about some kind of natural language interface, you don't necessarily need to be developing or fine-tuning LLMs yourself. You can be using an open source LLM, like today at the time of recording with you and me here, LLAMA 3.1 was released, their 400 billion plus parameter model is supposed to be competitive with the big foundation models like Claude 3.5 Sonnet and GPT-4.0.
- 00:33:21 But yeah, obviously that kind of huge model is going to be extremely expensive to run on your own infrastructure. So you're probably, for most use cases, you're going to be better off calling a proprietary API anyway, because it's going to be so much cheaper. But the point that I'm drifting away from here is that GenAI, like you're saying, it's possible that all of this attention on GenAI will mean that all aspects of AI will be elevated, but it risks, like you said, drowning out all the other kinds of data automation. And there's so many different kinds of approaches that could be appropriate in a particular circumstance, Bayesian statistical approaches, all the kinds of Frequentist modeling that we could be doing, simple regression models. There's so many ways that very simple data modeling could be exactly what you need, but yeah, you end up hiring people or pushing projects that are getting GenAI into products when ultimately users aren't going to use them. And then there's disappointment around, "Oh, we spent all this money on getting all these

GenAI features in, no one's using them. AI isn't that great for us anyway."

Nick Elprin: 00:34:31 Right. Yeah, I agree with you that especially the product integrations I think have a little bit of this jump the shark feel. But in defense of GenAI, I do think there are very significant opportunities to improve general workforce productivity across a number of different roles. And that's exciting, and I understand why that's exciting and compelling to business leaders as well, especially in an era of higher interest rates and things like that. It's great to be able to make a bunch of internal efficiency improvements. But I guess I haven't seen yet the GenAI use cases that drive a step change, like an order of magnitude of value through unlocking some new innovation or radical transformation.

Jon Krohn: 00:35:29 Yeah. We were talking about citizen data scientists, citizen developers, more like yes, the kinds of visual tools, low-code or no-code tools could end up being helpful. I'm sure there's lots of scenarios where that is exactly what people need. Maybe it makes it easier for people to quickly see how a system is working. I can see lots of use cases for that low-code no-code kind of tool. But I think even more transformative is what you just described around the way that generative AI makes it so easy for me as somebody who doesn't know a lot about computer science or MLOps, being able to have some code and have Google Gemini in my CoLab Notebook be giving me advice on what I'm doing, or GitHub Copilot is probably the most popular example of this kind of tool, it's wild. I did an episode recently, episode number 798 on Claude 3.5 Sonnet.

00:36:29 So, people can check out particularly the YouTube version of that. I mean, the audio-only version works as well, but in the YouTube version you can see how in that tool, if you're a listener to this show you should be paying

the 20 bucks a month to be getting access to Claude, because it's unreal how you can be generating what they call artifacts alongside your conversation, including fully functioning web apps. So, in terms of being able to pair program with an expert on whatever you want to be pair programming on, there's never been a better time and it's just going to get easier and easier.

00:37:02 So yes, there are definitely big kinds of efficiency gains that can be made from these kinds of tools, as you said. I'm just adding some color onto that, Nick. Daniel Hume, sorry to interrupt, but my last point here, Daniel Hume, whom we mentioned at the onset of this episode that you and I were together with him on a panel at NYU a couple of months ago, on his episode with us, episode number 807, he had an interesting take on how you could potentially get that kind of 10X lift in an organization.

00:37:33 It would require a lot of upfront investment, but he was making the case that in terms of the AI capabilities we have them today, if you were able to engineer your data pipelines to be taking the vast majority of corporate conversations, so Slack conversations, emails, meeting notes, if everything was being well digested by large language models, the context windows are large enough and you'd probably need several tiers of going to summarization to summarization, there's lots of cleverness that you would need to do in terms of stitching these systems together. But at least hypothetically the LLMs are there that they could be processing this huge amount of data, be identifying signals, and then advising people in the company up and down the organization on strategic decisions they could be making, even things that might sound controversial would be theoretically possible, like recommending who should get a promotion relative to other people.

00:38:32 So, I think there's a huge amount of potential but there's lots of work to do. Just having these wild LLM capabilities, which truly are wild, I don't want to understate that, it's mind-blowing to me. It's crazy to me that we have these technologies today, and two years ago, three years ago, I wouldn't have imagined that we would, but there's still tremendous amount of human ingenuity and hard work required to stitch these things together and make something powerful.

Nick Elprin: 00:38:57 It's possible I don't have a long enough foresight here and I'm not being imaginative enough, but I'm skeptical that the idea of we'll just ingest enough of the corporate data and the LLMs will be able to make good strategic recommendations, I'm skeptical of that idea, but we'll see. I mean, maybe the future will prove me wrong.

Jon Krohn: 00:39:20 Yeah. So, I can get your skepticism about just being able to feed a whole bunch of raw data, a whole bunch of raw natural language or other kinds of data, maybe even tabular data into an LLM and sort of auto-magically that cascades into business decisions. But you have on previous podcast appearances, as well as on the domino.ai website, your belief is that it is possible to weave AI into the fabric of businesses. So, how do you envision AI platforms integrating more closely with business strategies and executive decision-making processes today or in the future?

Nick Elprin: 00:39:57 Well, let me answer in two parts before getting to the platforms question. Let me start before we even talk about technology, and then I'll come back to the platforms part of it. So, what I mean by weaving AI into the fabric of the business is that the way for a company to get the most value from data science ML AI is to start with a really deep understanding of how their business works and what they're trying to optimize for and solve for, and then work backwards from there to figure out,

"Well, how can models or how can AI techniques help us do that better?" And that may sound obvious, but it's surprisingly uncommon. Far more common, unfortunately, is the reverse or the backwards approach where companies think like, "Oh, what can AI do?" Or it's like, "Well, what data do we have? And can we throw some AI at that and see what comes out?" And that's all backwards.

00:41:01 One of our customers is Verizon, and the woman who used to run data science there, I think she's chief data and analytics officer, her name's Linda Avery, she's given a great talk about this idea of working with your organization to build a value tree for data science and AI, which is to say, "Okay, what are the most important metrics to your business?" Forget data science for a second. And then what are the highest levers that drive those metrics, and what are your key strategic assets that you have as a business? And from there, you get down to, "Okay, well what are the creative or imaginative ways we can come up with for using data science and AI to move those needles?" So if it's an insurance company it's, "Okay, how do we better assess risk? How do we better price our policies? How do we provide a differentiated customer experience?"

00:42:01 Because if you believe that over time the competitiveness in the insurance space is going to be a race to the bottom on price, and if your strategy is win on customer experience, that's how you start. So then you say, "Okay, well how can we use data science to, for example, improve all of our touch points with customers and give customers the best possible customer experience?" But if you're an asset manager, then you're going to be looking at, "Okay, well how do we identify sources of alpha more rapidly? How do we build more automated investing strategies?" And that's the key lever to pull for your business. If you're a pharmaceutical company, it is, "How do we get

innovative therapies to market faster?" And then you work backwards from there. "Okay, well where's the bottleneck in our pipeline? Is it early ideation and discovery? Is it the clinical trial process and is that where we think we can shave time off the pipeline?"

00:43:05 And so, you take this top-down optimization problem by looking at your business and what are the key drivers you're trying to move the needle on? So, that's what I mean when I talk about weaving AI into the fabric of the business, it's not a thing you go and sprinkle on, it's a thing you say, "Well, as an executive team, we're planning our business. What's most important? What are the KPIs? What's the strategy? What do we view as our most important assets?" And that's how you prioritize your AI projects. Does that make sense?

Jon Krohn: 00:43:39 Mathematics forms the core of data science and machine learning. And now, with my Mathematical Foundations of Machine Learning course, you can get a firm grasp of that math, particularly the essential linear algebra and calculus. You can get all the lectures for free on my YouTube channel, but if you don't mind paying a typically small amount for the Udemy version, you get everything from YouTube plus fully worked solutions to exercises and an official course completion certificate. As countless guests on the show have emphasized, to be the best data scientist you can be, you've got to know the underlying math. So, check out the links to my Mathematical Foundations of Machine Learning course in the show notes, or at jonkrohn.com/udemy. That's jonkrohn.com/U-D-E-M-Y.

00:44:23 Makes a huge amount of sense, and I think you articulated it extremely well. But this is something that we've talked about on the show before, you're not the first person who's had this general hypothesis, if I'm reiterating this correctly, that you don't want to go from

seeing some new AI hammer, to trying to find some nails in the business to apply that to. You identify where are the biggest opportunities for automation? Where can we make the biggest strategic lift? And then as part of that solution, maybe AI can be one of the answers, maybe not.

Nick Elprin: 00:44:57 Right. And back to our earlier conversation, maybe GenAI isn't the right strategy or tactic for a particular opportunity. Maybe traditional ML is, maybe traditional statistics is. So, you work backwards from what's the metric we're trying to move, what's the best way that AI broadly can help with that, and then what's the best technique to apply or tactic to apply? And notice, when you asked about how I see AI platforms fitting in, notice nothing I've said so far has anything to do with technology or with platforms. I mean, I think the idea that, oh, you just bring in some tool or some platform, and all of a sudden that you're going to be this AI driven, model driven company, that's not how it works. It has to start with people, and process, and technology. Don't get me wrong, technology and platforms can be a very valuable accelerator of that change, but in and of themselves, they won't drive that change.

00:46:02 So, they're an enabling component once as an organization you know what you're doing and what you want your strategy to be. So then I think about, "Okay, well how can platforms like ours help?" And it comes back to what we talked about in the opening part of the conversation. I mean, so the first thing is if you want to weave AI into the fabric of your business, inevitably you're going to have different use cases calling for different techniques and different tactics. Those are all going to need different types of modeling projects, they're going to use different types of tools. And so, a good platform will give an organization the agility and the flexibility to quickly and easily use the best tool for the job. "This is best addressed with regression, how do we do that? This



is best address addressed with GenAI, how do we do that? This is a deep learning computer vision problem, how do we do that?"

- 00:47:03 So, a good platform will not box the company into one particular methodology. It's not like, "Oh, we're going to do everything with auto ML." It's how do we give you the most expansive toolbox? Because different jobs call for different tools. That's the first thing. Platforms can make it easier and faster to compress the time from idea to production impact. And that's another pitfall for a lot of organizations is like, "Hey, we did a lot of prototyping. We've got some exciting projects, but okay, have they delivered business impact? Have they touched your business processes where they're affecting real life decision making, either through automation or by informing a human? Not everything has to be fully automated, but are they, in finance we'd say are they touching the money? And there's an analog in other industries.
- 00:48:15 Are they affecting decisions or actions? And that's something a platform can help with by streamlining the process to get models reviewed, packaged into forms that are more conducive to production operations, handling the operations of them, doing things like model monitoring so you can mitigate risk and be confident these things are deployed in a safe environment. So, that's a lot of what we think about. And then the third thing I would say in terms of how platforms can help is again, back to my point that technology is almost never the end in of itself, it's a means to an end. And the big thing, huge thing companies have to think about as they weave models into more parts of their business is scaling their human organizations. You need more data scientists, you need to go from 10 to hundreds.

00:49:15 And so, how do you attract and retain talent, and how do you upskill talent you have, and how do you make your talent most productive? Those are things that platforms play a huge impact in. A big driver of why top talent leaves organizations is they don't have good productivity tools. They don't like the way they have to work, it's too annoying, it's too painful, it's too cumbersome. They have to wait weeks to get machines from IT. That's really frustrating for great data scientists. So, how do you give your data scientists a great experience so they enjoy working? Some of the companies we work with have said that in their recruiting process, they have candidates use our platform and they say, "Look, this is what you'd be using if you came to work here." And that's a selling point, that gets data scientists excited.

00:50:14 And then back to this other thing we talked about early on of this idea of preserving organizational knowledge and making it so folks don't have to continuously reinvent the wheel. It's a massive productivity drain if every time you hire more people or someone leaves your company, you lose institutional knowledge. And new folks who are joining take three months to ramp up because they're reinventing a bunch of stuff. So as you scale your data science organization, which is critical to delivering more models, having more throughput, addressing more use cases in your business, how do you make your talent more scalable? So when I talk about scalability, I'm not talking about the computers. It's not like, "How many GPUs are you running?" I mean, how can you scale your data science organization efficiently? And that's another big role that a platform can play.

Jon Krohn: 00:51:12 When does a company realize that they need a unified AI platform like this? How big does a company need to get? You must have this kind of experience where you talked about how Domino, for example, is best suited or can make the biggest impact with large enterprises. When do



you realize that you need to make that jump, that it's worth the investment to integrating a tool like this into your business?

- Nick Elprin: 00:51:39 This is going to sound like a weirdly precise answer, but in our experience it's around 20 people, and there's no-
- Jon Krohn: 00:51:46 Like 20 data scientists or 20 people?
- Nick Elprin: 00:51:48 Sorry, 20 data scientists. Yeah. Which to be clear, you can have that size of a data science organization in a 500 person company or in a 10,000 person company, but it's around a couple dozen people doing data science, computational research work. And that's an empirical answer, not a first principles derived answer, but in our experience that just tends to be where somebody ... At that point, you have someone who's responsible for the data scientist who is looking at all from an altitude and thinking with a certain level of conceptual perspective to think about, "How do I make this work better, and how is my machine working?"
- 00:52:38 As opposed to just having strong data scientists running around, so you have someone who's thinking about the organization. And then you're at a scale where you start to have people duplicating work and you start to notice a lot of inconsistency, or everyone's got their own AMI that they've set up for their AWS usage for how they like to do their work, and they all have different packages installed. And now you're noticing it's hard for people to share work. And that's when this need for, "Hey, we need some standardization, we need some consistency," that's when it kicks in.
- Jon Krohn: 00:53:13 Nice, that makes a lot of sense. And it's nice to have a weirdly specific number, like 20, to work with that benchmark. Gives us a great soundbite too for Eclipse. Nice. So back in 2016, in an event hosted by Matt Turck

of FirstMark Capital, not First Matt Capital. So yeah, Matt Turk of FirstMark Capital, he's a really well-known individual in the AI space, particularly at the intersection of AI and VC around fostering and classifying tech companies in the ecosystem. And so, back in that event in 2016, eight years ago, regarding senior level management thinking about silver bullet technologies, you said, "Technology itself is not a solution, and that people, process and technology must work together." So, in the past eight years there's been huge transformations in AI and data technologies, generative AI being the latest example. So, there's been a lot of maturity in the past eight years. Do you think that people and processes have kept up with what the tech can do, and do you have guidance on what could support a culture shift?

- Nick Elprin: 00:54:30 Gosh, it's quite the throwback there. Well, and first of all we should all appreciate Matt Turk's prescience. He started building the, what did you call it? What's it called? This data and AI landscape chart back almost 10 years ago. And back then it was just a couple products in that whole thing. And now, I don't know, you got to print it out on a scale to fill a wall to be able to see it. So yeah, he was certainly ahead of the curve in spotting where all this was going. So, I say all that just to stall as I think about an answer to your question.
- 00:55:09 Well first, I continue to strongly believe what I said, that no piece of technology is a panacea or a silver bullet. And I think if anything, the last eight years have demonstrated that in how rapidly the state of the art has changed. With the state-of-the-art LMs, there's a new one every three months now. But like you said earlier, even three years ago only in small circles were people talking about generative AI. So, we went from data science, to deep learning, to generative AI, and like I said, anyone who went all in on AutoML five years ago is now realizing, "Hey, actually that's not everything that we need to do."

But you look at some of the other technologies, I mean, this is going back before eight years, but we can probably all remember when Hadoop was the thing for doing large scale data queries and analytics, and okay, that got blown away by Spark. And now I think Spark is a lot less performant than Ray, and there'll be some new successor to that in another couple of years. So, my point at the time was there's always a next thing.

00:56:34 And anyway, then now back to the other part of your question about how have people in process kept up, to generate business impact from models, or from AI or data science, it is not merely a question of can I build the model? It is also a question of can I ... Well, first of all, it's a question of can my organization work in a way that prioritizes the right things for us to build? So, this is back to what we talked about a minute ago of what I mean by weaving AI into the fabric of your business, do we have the right cross-functional collaboration for data science organizations to get the right priorities to know what's most important to work on? So, it's the leadership team working together to prioritize use cases in the right way that has nothing to do with technology.

00:57:24 Then let's say we build a great model, do we have the organizational capability to get it into production in a way that is safe and streamlined? There are companies we've worked with that have said it has taken them longer to get a model into production than it has taken to build a new headquarters, and that's crazy. I know I've been avoiding directly answering your question about whether this is all kept up. I think the way companies think about evolving, maturing their internal processes and organizational capabilities was evolving in a good way. We were on a good track. I think GenAI is a bit of a disruptor in that, because let's think about model approvals, model reviews, model governance, model validation. GenAI introduces a whole new set of complexities for how

organizations need to review, and govern, and manage models that are in production.

00:58:36 I think companies were getting comfortable with model drift. They were getting comfortable with, "Hey, let's review models before they go to production to understand feature weightings, or other explainability analysis." Okay, but now you try to do that with, "We're going to put a GenAI model in production," it's like, well, now we got to invent whole new processes for, well, what are our standards for guardrails? Do we have any idea what data was used to train this? How do we monitor these things now they're in production, because it's not quantitative outputs you're getting? It's hard to precisely define what drift would mean.

00:59:19 And like you pointed out with your example of the WHOOP chatbot, we're still in a bit of a hype bubble around just throw chatbots at everything instead of being disciplined to think through what are the best applications for GenAI to really move the needle for our business? So it was a very, very rambling answer, but I think my summary would be two years ago I think companies processes and internal discipline for maturing how they used AI was going in a good direction, and now we've had to reset a little bit because of the differences and the nuances of GenAI.

Jon Krohn: 01:00:06 Yeah, that answer makes a lot of sense, and that's a great answer. So, switching gears a bit here. From the way that Domino can fit, or tools like Domino can fit into business processes and allow enterprises to automate things, to have their data science team scale more readily to ultimately get models more rapidly from prototyping into production, I'd like to talk a little bit if you can, and there may not be a lot you can go into, I'm not sure, about the way that this works, like technical specifics behind the hood. Yeah, it'd be really interesting to hear how if you

are Domino, if you're creating a tool like yours that data scientists, other technical people that whole large enterprises are going to be relying on, what do you do when you build your tech stack to ensure that it is highly scalable and performant?

- Nick Elprin: 01:01:00 Oh, I thought you were maybe going in a slightly different direction. So, less about how the product works and more how we ensure scalability? Or maybe-
- Jon Krohn: 01:01:10 You can do a little bit of both if you want.
- Nick Elprin: 01:01:12 Or would you mind giving me a slightly more specific question?
- Jon Krohn: 01:01:19 Sure. So, what kinds of programming languages are in the stack to create Domino Data Lab?
- Nick Elprin: 01:01:26 Yeah, yeah. Sure. So, most of our platform is built in Scala, and that was a choice we made very, very early on and we've stuck with. Our front end, as you'd imagine, is JavaScript, and we use React for that. I think we initially had a different front end framework but have switched to React. And then deep in the stack, we're a fully Kubernetes native platform. And so, we do have some bits of the code base that are directly written in Go because of some kind of direct interfacing we need to do with the Kubernetes stack. We've built some custom Kubernetes operators and things like that.
- Jon Krohn: 01:02:16 Very cool. Yeah, that makes a lot of sense. Yep. React is a front end, not hugely surprising. It's also for people who are building platforms, they're probably aware of how something like Kubernetes is critical to allowing these platforms to be highly scalable. I hadn't heard of people writing their own Go code to make aspects of that Kubernetes customizable, so that's cool and new to me. And then yeah, interesting to hear the platform being built in Scala from the start as opposed to say something

like Python, which might've been another option for the backend at the time that you would've started on Domino Data Labs. So, I don't know a huge amount about Scala. It's a functional programming language, right? Am I right?

- Nick Elprin: 01:03:00 It is. I mean, used in its purest form it is functional. It's got a lot of flexibility, and it's also possible to use it more imperatively. And some Scala code looks a lot like just more concise Java code. It's a good hybrid.
- Jon Krohn: 01:03:19 Cool. Well, it's great to have that insight into what's going on under the hood. It's the kind of question I should probably ask most of our guests on the show, particularly technically minded founders like you are. And speaking of founding, yeah, go ahead.
- Nick Elprin: 01:03:35 Well in full disclosure, I haven't touched a line of code in probably six years or something. And pretty early on I stopped being allowed from writing too much code because I wasn't a good enough engineer.
- Jon Krohn: 01:03:48 No, I understand that completely. I'm in exactly the same boat, but you do have a computer science masters from Harvard, and you spent seven years as a senior application technologist at Bridgewater. So, that's what I mean by technically minded. I don't necessarily mean pushing production code, but you have awareness of these kinds of things, no doubt, and these kinds of big strategic technology decisions that get made in your company. So speaking of that journey a little bit, what caused you to make that jump? Because if I am reading this correctly, and I'm just going off your LinkedIn profile so you know more than me, but it looks like from there that you went from doing undergrad at Harvard and Computer Science, Masters at Harvard and Computer Science right into working at Bridgewater Associates for that seven years, and working as a technologist there.



And then it looks like about a year after leaving Bridgewater, you co-founded Domino Data Labs and had been CEO of that company since, which is coming on a dozen years now.

01:04:52 And so that could be, if that information is complete information, that is making that big leap from employee, safe W-2 employee at a super successful, globally renowned brand, and then making that jump to co-founding a company, you couldn't have expected that you'd have the success that you would have. There's always that hope as a founder, but you're in the top percentile kind of thing of tech co-founders and the success that you've had. So, what drove you in that initial instance to make that big leap? And was it scary, or has it always felt just like the right thing?

Nick Elprin: 01:05:38 Yeah, I'm trying to put myself back in the head space I was in back then. Well, so what drove me initially was really two things, and one was more extrinsic and one was more intrinsic. The extrinsic one is, I mean, it's so cliché, but I wanted to feel like I was having a bigger impact on the world. Rather than being at one company and helping that company, I really wanted to do something that I felt would help more organizations more at a bigger scale. And the second thing, which is a little more internal was I need to feel like I am constantly learning, and growing, and improving. And if I don't feel like that, people talk about being in the stretch zone, this feeling of you're almost drowning but you've just got your head above water, but then you finally figure it out and then you're onto the next challenge, I needed to feel that.

01:06:49 And I got to a point where I felt like my own personal growth was stagnating a little bit. And in retrospect, yeah, I think it was scary to make the leap. And as a result of being scared to make the leap, I probably waited a little longer than I should have. I probably put off making the

leap for, I don't know, maybe a year after I really knew I wanted to do it. But what finally pushed me to do it was this fear of future regret. Like all human beings, I felt myself getting older. As I was getting older, I felt myself getting more risk averse, not less. And so, I had this moment of projecting my life 10 years forward and thinking, "Well gosh, if I didn't make a change now I might end up in a situation where I was so comfortable and so used to what I've been doing that I would've just been stuck there forever."

01:07:58 So, that was the moment where I was like, "Okay, well there's never going to be a better time to do something different or try something different." Yeah, I mean, in the early days, I have two co-founders and we'd all work together at Bridgewater, and like you said, we didn't know at the time how big what we were doing was going to be. We were just very focused on making product that data scientists would love to use, and that was our North Star. And so, we spent a ton of time sitting with data scientists, working with them, and we were focused on what's going to make your life better? And that guided how we iterated on the product in the early days, and then I think we got somewhat lucky with the timing of how data science took off and then how it turned into ML, and now of course AI. So, always [inaudible 01:09:00]

Jon Krohn: 01:08:59 Yeah, congrats. Do you have one tip for our listeners on how they might be able to recapitulate the kind of success you've had? If they're looking to create their own startup or maybe recently have, what can they do to try to have the kind of tremendous success that Domino's had?

Nick Elprin: 01:09:13 Well, I think it's cliché, I'm certainly not the only one to say this, but for me it really is about just focus on customers, focus on creating value for the customer. If you put the customer first, then the rest of the good stuff will follow.



- Jon Krohn: 01:09:26 It is a cliché, but a viable one. The other one that I thought you might go for once you started saying that sentence was it's all about building the right team.
- Nick Elprin: 01:09:34 That happens next, in my experience. That's the next phase.
- Jon Krohn: 01:09:39 Yeah, put the customer first, love it. So, awesome. Thank you for this amazing interview, Nick. Before I let you go, I always ask my guests for a book recommendation. Do you have one for us today?
- Nick Elprin: 01:09:51 Yeah. Well, it's been a long time since I've read it but it's one of my all time favorite books. It's called The Lessons of History, and it's nonfiction, and the story behind it, which it's amazing so I'll just tell the quick story. It's written by a husband and wife historian pair, Will and Ariel Durant, I think, are their names. And they spent their whole career teaching history, and they wrote a multi-volume encyclopedia of the entirety of mostly Western civilization history, and it's like eight volumes and thousands of pages. And then as they got toward the end of their career, I think they were in their 70s, they wrote this book called The Lessons of History. It's 100 pages, and it's like their synthesis of everything they have learned and ever written boiled down to just the essence of these core recurring patterns that they noticed across centuries, and different civilizations. And the signal-to-noise ratio in the book is unlike anything I've ever seen, it's just every word is rich. So yeah, that'd be my recommendation.
- Jon Krohn: 01:11:12 Nice, very cool. And you've been an amazing guest, we've learned a ton from you. How can listeners follow you and hear your thoughts after this episode?
- Nick Elprin: 01:11:23 You should follow the Domino Data Lab page on LinkedIn. I personally don't have a lot of social media



presence, but so LinkedIn for myself, Nick.Elprin or Nick Elprin on LinkedIn, or the Domino company page.

- Jon Krohn: 01:11:39 Nice. Thanks a lot, Nick. And yeah, hopefully catch you at some part of the world soon, New York, San Francisco, wherever, and really appreciate you taking the time with us today as the CEO of such a tremendously successful company. We are grateful to be able to have that time with you, and yeah, be able to glean some rich insights from you.
- Nick Elprin: 01:11:59 It was a real pleasure. Thanks, Jon.
- Jon Krohn: 01:12:05 What a great episode. In it, Nick filled us in on how at around 20 data scientists it becomes worthwhile to integrate a comprehensive platform like Domino for mission-critical AI, how starting with business problems and devising AI solutions only where necessary is more likely to lead to commercial success than trying to apply new AI hammers to any nail you can find, how using Scala instead of Python, as well as Kubernetes with customized Go components for their backend, has allowed the Domino platform to be so performant. And how putting the customer first has allowed Nick to develop a wildly successful data science startup on his first attempt.
- 01:12:43 As always, you can get all the show notes, including the transcript for this episode, the video recording, any materials mentioned on the show, and URLs for Nick's social media profiles as well as my own at superdatascience.com/811, 8-1-1. Thanks of course, to everyone on this Super Data Science podcast team, our podcast manager, Ivana Zibert, media editor Mario Pombo, operations manager Natalie Ziajski, researcher Serg Masis, writers Dr. Zara Karschay and Silvia Ogweng, and founder Kirill Eremenko.



- 01:13:16 Thanks to all of them for producing another exquisite episode for us today for enabling that super team to create this free podcast for you. We're super grateful to our sponsors, and you can support the show by checking out our sponsors links, which are in the show notes, so please consider doing that. And if you are interested in sponsoring an episode yourself, you can get the details on how by making your way to jonkrohn.com/podcast.
- 01:13:41 Otherwise, please share, review, subscribe and all that good stuff. But most importantly, just keep on tuning in. I'm so grateful to have you listening, and hope I can continue to make episodes you love for years and years to come. Till next time, keep on rocking it out there and I'm looking forward to enjoying another round of Super Data Science podcast with you very soon.