All right, hey everyone, welcome back to Coding Temple's Agile Leaders

We help product and tech leaders execute and innovate in these rapidly changing times with more confidence and velocity

I'm your host, Evan Shye

Today we are joined by my friend, Greg Lachnane, the co-founder and CEO of AI Maker Space, where him and his incredible team are building the world's leading community for building production LLM applications

So this episode is for the technologists out there looking to push the bleeding edge of AI in their organization

Thank you so much for being here, Greg

Yeah, thanks for having me, Evan

It's awesome to be here with you

Well, I have the pleasure of knowing a little bit about your background and how impressive it is and why you are my go-to source for everything related to AI, but maybe you can share a little bit about your journey with the audience here

Yeah, you know, in 2015, I graduated with my PhD

It was an optimization at the time

And part of the reason I even got it was because I kind of loved teaching

And I started doing that as an undergrad

So quickly kind of came out of school, got into government contracting, consulting and started work with teams to commercialize tech

And that kind of led me into liking the startup vibe more than the big companies

And I found it was very hard to make money with research

And even though I started to do some consulting outside of government contracting with actual startups, I found it was really still hard to make money with AI at the time, 16, 17, 18

And I started doing a lot of sort of just research for the CTO and you're kind of taking this work that you do and you're thrown it over the wall of the engineering team and it often dies on the mind

You know, this kind of pushed me more and more towards trying to get inside and build something with a team towards product management

So I kind of went and I went to entrepreneurship school, dabbled in some startups that weren't mind all along I'm teaching and then COVID hits

And I realized like, man, online teaching and learning is like, this is like the future

I like, I have got to get into this

I've always loved teaching

It's never really paid, but I got the feeling at the time this is what I should be focused on

2020 actually take a full time job teaching

So I'm teaching five days a week

I'm trying to become like the best at online teaching and learning

And you know, it was a big opportunity because still most people have never been to a Zoom room

That's like really, really well run

It's not like a thing that people have experienced

But it can be done

And you know, this is something I've been working on for the last four or five years

I tried to launch a course back then teaching people 3D printing

It was a disaster

I don't really know what I was doing

Headed out to Silicon Valley to sort of learn startups, learn the industry of EdTech, understand boot camps

It's kind of got me into a role at the company called fourth brain, which is very close to deep learning AI and got me kind of close to a lot of people that are kind of at the top of the industry

I looked around and I realized like this is what I want to be doing building these boot camps, but this sort of VC product, VC backed, EdTech boot camp business model

It's not quite right

And so, you know, after Chad GBT came out, 2023, I realized, man, this is the year I want to do this a little bit better than I've ever seen anybody do it

This kind of led to what AI maker space has become today

So we're building boot camps out on the edge of what's possible with open source

And that's LLM today in production

I love it, man

I knew there were some similarities in our journey, but I didn't quite appreciate how much

I guess I've been an entrepreneur my entire career, but did some time in an academia, a few of us in a PhD program and was publishing research and realizing how far of a distance or golf there was between the work we were publishing and the impact I was having on people's lives and ended up starting trying to do skill development, building applications to deliver to that skill development, to make a bigger impact in people's lives

And then, you know, a lot of philosophical alignment in terms of how can we do that better and ultimately brought both of us here, I think

Were there any skills that were kind of most transferable from that past life into more of the operator seat, product leader seat? Yeah, I mean, you know, I think when I look across all of it, it's kind of like this willingness to sort of like attack one thing as hard as you can at each period

So it kind of is like resilience is number one

It's kind of like cookie cutter thing, but like, you know, like I was deeply under employed for long stretches, you know, throughout this journey

And it's just part of like continuing to just smash the grind

But each sort of skill I learned a long way, like doing research, very important if you're going to do consult things, right? And then building a startup, actually the way you build it out to begin with and find product market fit and figure out what you should be building is kind of through consulting

And so like, you know, each of these things kind of stacks on each other

And then it was very, very helpful to go sort of get the broad perspective of entrepreneurship school, dabble a little bit of this and that

And then as a CEO, I'm finding I go back to some of that, you know, okay, I have a CFO day

I got to go pick up how to actually do the balance sheet stuff

Okay, I have this kind of day

And, you know, I think that that's been, it's just kind of all come together because the startup CEO is kind of an everything job

So you kind of have to stack it one piece at a time

Amen, man

Even something as seemingly singularly constrained to science or research, but just the scientific method or running experiments appropriately and developing hypotheses

And it always doesn't have to be this like perfectly elaborate experiment, but even just thinking in ways where you can actually test hypotheses or expose where you're wrong

I think is a useful thing for a lot of entrepreneurs who are trying to

That's right

And I think like I've talked to, you know, mentors of mine, whatever, it's like this idea that what is your hypothesis? Because the thing about business is like, you have to figure out what's true about your customers, you have to figure out what's true about what's going wrong in your company

And it's like, it's like the opposite of sort of like getting emotional about it and whatever you feel about it

And so you've got to be able to say, okay, and the decision making framework of the hypothesis is very useful because otherwise like if you're trying to think about everything and how it engages in interacts, like you can't make any decisions

It's like this hypothesis, this is how I'm going to test it

You know, it's very, very, it's exactly that in business

Yeah, that you have to have a scientific method, you know, let's get to the truth as fast as possible because we're going to get there one way or another and like that's not pay a toll for the fee tour, you know

So kind of leaning back on, I guess some of your consulting work and then now what you're seeing with technologists and technical leaders inside enterprises, you know, obviously a lot of organizations, most probably organizations are starting to really look at AI and recognize that there are meaningful implications that this is going to have on our business

But what do we ultimately do about it? What is the strategy? How do we actualize that strategy and create value for the organization? Is there a common thread or a primary set of challenges that you see technology leaders face when implementing AI strategies in their organization? Yeah, I mean, I think that you have to sort of break down the industry into its component parts and you have to break down the technology at the same time

So, you know, in industry, you're going to prototype things and then you're going to put them into production

On the technology side, we really want to kind of differentiate between AI sort of classical machine learning, deep learning and then this sort of generative AI that we're

talking about when we talk about large language models

Of course, there are many, many fine lines there, but from a what you should be building for your customers perspective, you know, you want to understand, okay, top level, yes, I need to go from prototype to production

Yes, I need to understand AI versus generative AI

But you know what your customers want and what is going to make their lives easier

And like part of the thing is, part of the dirty little secret is like, well, most products today that are built do not actually require AI

And you know, of the ones that do require AI, most of them do not require generative AI

So you end up when you go and you take it, I'm only going to create business value from my customers perspective

You end up kind of taking this unsexy approach of saying, well, maybe I'll just do sort of classic product management

Maybe I'll do more classical AI product management

But you know, the board's asking about LLAMs

Everybody wants to know what we're doing with LLAMs

Chat GBT came out

Everybody's freaking out

What are we doing about it? It's important to have something there

And it's important to start moving down this prototyping to production pathway

And so this is where I'll sort of leave the classical AI aside and talk about generative because that's what people want to hear about today

I think one of the things is there are companies that are good at prototyping already

They're great at it

And they're sitting there and they're like, we got these prototypes

We're dangling the candy in front of people that are in the company

They're like, man, what are we waiting for? Let's push these things to production and scale them out

These are amazing applications

And there's a real hang up there in production

So these are the financial industries, this is the insurance industry

This is these are the guys that have an all digital product

So it's all digital

It's all about the data

It always has been, you know, quants are the OG, ML and AI people

Making money off and making money has always kind of led the way here

So these industries that build products based on that idea are still very much leading the way

And so, you know, when I talk to folks from industries like that, what you end up hearing is going to appear

Well, a lot of these challenges of going to production, the infrastructure for LLMs in production is different than the infrastructure for AI's in production

It's classical AI in production

And by the nature of the large in large language models

And so you end up like having to figure out how to get new tools

And there are a lot of these open source edge tools that, you know, they're not well supported

So in production, right, if something breaks at 1, 2, 3, 4 in the morning, who are you going to call to make sure that you can fix it? And if the answer is, I don't have anybody to call, well, we don't want to go to production

And this is a very important gap right now in the industry that we're going to continue to try to create some value for folks in that space is one of, is one of our key strategic goals of this year

So because we believe that's the open source edge

Now, most people are not there, all right? Most people are back in the productions in the prototyping phase

And in the prototyping phase, you have to sort of decide what should I be building? Why? How am I going to leverage tools like prompt engineering, like fine tuning, like retrieval augmented generation, bragg? If you can't sort of intelligently even speak about these topics, let alone how they relate to one another and how they come together holistically to create an app that has human level performance

That's aligned with your customers

That's giving a great experience

That's not going to get you in trouble

And you're not very good at prototyping yet

And you should be spending your energies and your focus there

So I think what you have to do is you have to look across your organization

You have to be real honest about how mature are we from a technological standpoint

Some industries are just less mature, manufacturing, defense

They tend to move slower and that's OK

Some other industries feel the heat of competition, the insurance, right? The financial, they're like, if we don't get on this, we're out of business

So if you have to look at your industry, you have to look at your actual executive team and the capabilities you have from a data science and machine learning perspective now

And you have to start building on that wherever you are

And whatever that looks like is going to be different for your particular situation

But it always starts with proof concepts and prototypes

It always moves into production

And you're always going to have to make that differentiation between classical machine learning deep learning and generative AI and no AI at all

So it's tough to answer the question in general

But that's my best right there

It's a very, very useful framing and mental model that helps kind of bifurcate and organize the different contexts that businesses or maybe the different phases that businesses are ultimately in

And they're interesting points in both elements

I guess let's start with the

Well, actually, do you have a general sense or pulse, at least even anecdotally just based on your experience? Where do the majority of businesses fall right now? In the production phase or the prototyping phase? Bass majority of businesses are stuck in prototyping

Very, very few businesses that I've spoken to are actually out like

They're like, yeah, yeah, we get it

We understand how to use all the latest and greatest tools

We're building that

Our people understand it

The thing is we have these sort of like

Other infrastructure, security, whatever issues that are halting me now from creating massive value with these tools tomorrow

Right? And so that's the edge

Most people are somewhere getting started

One, two, three POCs, prototypes

You really have to sort of maybe get five or ten that actually really worked and you know you want to go to production with

So it's kind of a long road from starting POCs to being ready to rock

Totally, totally get that

We'll spend a meaningful amount of time in that kind of earlier phase and how you go about approaching it to hopefully create as much downstream value for customers as possible

Although I will say that you really nicely elucidated some seemingly inevitable challenges that companies will face on the production side of it, particularly as it relates to skill gaps and capacity planning and all of that to ensure that they're not only able to launch it in production but maintain it and continue to iterate on it

Yeah, yeah

The iteration has a whole other piece

Yeah, yeah, yeah

I'm glad I bet

Okay, so we're in the earlier part of the journey as a company, you know, really evaluating

Okay, we know, well, we have some early hypotheses about how we can create value for customers

I guess are there the frameworks with which you would advise product leaders or technical leaders to utilize out of the gate to ensure that they are consistently orienting around customer value and not kind of getting distracted by shiny objects? Yeah, I mean, I mean, I would just, I would start with, you know, your last quarter, your last year's strategy and, you know, it's like, well, what are some of the initiatives that you have on there that could potentially be augmented with some of these newer technologies? You know, like that, that, that sort of easiest, lowest hanging fruit

And then one of the other things is, I mean, that I recommend to people is like, go around to, you know, your, your functional department head, your functions, your departments, whatever

You, you are, and you ask people to leave them

It's like, what are some things that your people like complain about having to do all the time that are pretty monotonous? Like, what are the things that, you know, and you're often like, it's a customer support, right? It's a, it's the same question I'm answering over and over

It's the same freaking pipe I'm fixing 10,000 times

And the different maintenance person has to figure it out exactly the same way, even though it's already been done 15 times before

It's like, where are you not transferring that sort of, that sort of learned experience and knowledge to the next person, the next person, the next person? And how can you sort of automate that? Or what questions is HR getting on, you know, policies constantly, right? You sort of question answering systems or great sort of framework and mental model of use

Like, where are people asking questions that you get the same answer over and over, right? And, and, you know, that, that to me is probably the place to start is, what's bothering people about what they have to spend time on? That's sort of an internal application thing

So questions are you answering for customers over and over and over? That's sort of the number one external application thing

And then, you know, look in your own strategy, maybe you can start there and see what you can come up with

I love that

The, the internal focus first feels like to your point, lowering fruit, safer, more reliable

I mean, by definition, you can talk directly to the customer, the customer is part of your organization

So probably a safer space to be

You get, you capture some fast wins and you start to develop more confidence in the muscle and in the technology and you could probably build on that to some degree

Yeah

Yeah

And, and I think like one of the classic ones, even from classic A

I

was like the auto quoting sales tool, right? It's like, why are we, why are we still manually quoting people? Like, but can you actually build something that your sales team wants to use is the real question, right? Not that you're like forcing them by hooking cropped to sort of like, no, you have to pick this up otherwise

Yeah

And so are you actually building things that people want? It's as much either testing site too

It always matters, man

Is this something people actually want? And the best way to find out is to go talk to them

It's amazing how challenging that can be sometimes though

Before we dive into some more specifics around, you know, data-centric approaches and whatnot

Overall, do you have a sense of where you think there are real competitive advantages in A
I
for businesses? I sense that, you know, maybe to some degree, it looks very similar to how it's historically looked or where advantages or value opportunities already exist with customers
Is that kind of your position? Are there nuances to that? Yeah, I mean, like one use case that jumps out to me today for enterprise for large businesses is, you know, one thing we like to do is we like to go and do events for people and we'll go and we'll open up their web page and we'll show that their web page chatbot is just absolutely garbage, right? We'll ask it a question
We'll be like, hey, can you answer this simple question of this policy that is clearly on your website and I can find on the other page? And it literally can't answer the question
And you're like, what's the point of this thing? And the answer is baked into whatever sort of website creation thing that you used
And that's sort of the stats of these chatbots
Yeah
So it always goes back to chatbots, right? So, you know, it's like the, what you can do today is you can build a simple retrieval augment generation system
You can web scrape your own website and you can build this thing independent of your website as a question answering AI that returns a coherent answer and also returns you specific links to exact reference information and documentation where you found it
Now, this is sort of level one
But what you want to think about is when your customer asks a question into a live chat, what's the person actually doing? They're coming in and they're saying, well, let me see, what exactly have we done with Evan before? Let me look and see what his entire, what products has he bought for us? How many times he's called support? What's Evan in his general demeanor towards us like? How should we be engaging with him? How could we potentially upsell or what could we potentially do right now to make his life a little easier or better? Like, there's a complex sort of meta problem amongst all possible Evan's and Greg's that people have ever had in their customer base, right? And the thing is like, if you don't go and you don't start building your own solution to this today, what's going to happen is