Welcome to the OpenAI podcast, the podcast that opens up the world of AI in a quick and concise manner.

Tune in daily to hear the latest news and breakthroughs in the rapidly evolving world of artificial intelligence.

If you've been following the podcast for a while, you'll know that over the last six months I've been working on a stealth AI startup.

Of the hundreds of projects I've covered, this is the one that I believe has the greatest potential.

So today I'm excited to announce AIBOX.

AIBOX is a no-code AI app building platform paired with the App Store for AI that lets you monetize your AI tools.

The platform lets you build apps by linking together AI models like chatGPT, mid-journey and 11 labs, eventually will integrate with software like Gmail, Trello and Salesforce so you can use AI to automate every function in your organization.

To get notified when we launch and be one of the first to build on the platform, you can join the wait list at AIBOX.AI, the link is in the show notes.

We are currently raising a seed round of funding.

If you're an investor that is focused on disruptive tech, I'd love to tell you more about the platform.

You can reach out to me at jaden at AIBOX.AI, I'll leave that email in the show notes.

So San Francisco based startup Kalina has recently announced a $15 million funding round led by Lobby Capital with additional investments from SignalFire and Bloomberg Beta.

So this infusion of capital now boosts the company's total funds to $21 million that they've raised and this is aimed primarily at scaling its research team and also forming alliances with regulatory bodies and amplifying sales and marketing initiatives.

If you listened to my recent interview that I did yesterday with Bradley from Project Voice, it was really interesting because I asked him, you know, like what are some things that you look for in startups that a lot of people are missing?

And he said, one of the biggest things was that they need startups today need to be at least thinking about regulatory frameworks, what's going to happen in the future.

And he's like, if you're not looking at that, you're really going to be left behind because right now there's so much regulation coming down the pipe, there's so much changing in the space.

And he's like, you know, a startup today could become completely irrelevant if the wrong regulation comes down the pipe.

So, you know, no one can predict exactly what's going to happen in that, but startups need to be aware and following and, you know, essentially let investors know that this is something that they're cognizant of because this is incredibly important.

So to me, I find that interesting when, you know, Kalina right here just raised $50 million and one of the big things that they're doing, right, this is essentially like their company is set to kind of tackle trust deficit in AI with advanced model testing frameworks.

So one of the biggest things they're putting their $50 million to is forming alliances with regulatory bodies.

So very, very interesting.

So Kalina was actually founded back in 2021 by Muhammad, Al-Jedi, Andrew Shee, Gordon Hart and the trio has a lot of extensive experience in AI departments across a bunch of different industries.

So Amazon, Palantir, Rokton and SNAPs and their venture essentially is looking to address a really fundamental issue in AI deployment.

So it's the absence of trust among both its developers and the public.

So this is what they said, quote, and this was actually Al-Jedi.

So the use cases for AI are enormous, but AI lacks trust from both builders and the public. This technology must be rolled out in a way that makes digital experiences better, not worse.

The genie isn't going back in the model, but as an industry, we can make sure we make the right wishes.

So I think kind of going beyond some of the, you know, conventional metrics.

I think what really kind of sets Kalina apart is its comprehensive, quote, and quote, model quality framework.

So this is a tool set designed to provide robust, customizable and enterprise friendly model testing solutions.

So unlike other platforms that concentrate solely on, you know, component level testing, Kalina offers an end to end testing of AI and machine learning products.

So Al-Jedi again said, quote, first and foremost, we want to provide a new framework for model quality, not just a tool that simplifies current approaches.

And he really was kind of emphasizing a couple of things, but I think really with Kalina's user interface, teams can construct test cases to rigorously assess a model's performance across multiple criteria.

So it also kind of highlights potential gaps in AI model test data coverage and tracks associated development risks.

So the platform aims to shift the focus from, you know, blanket metrics like accuracy scores to more nuanced evaluations.

You know, Al-Jedi elaborated, quote, for example, a model with a 95% accuracy in detecting cars isn't necessarily better than one with an 89% accuracy.

Each has their own strengths and weaknesses, detecting cars in varying weather conditions or occlusion levels, spotting a car's orientation, et cetera.

So I think AI engineers reportedly spent around 20% of their time on analyzing and developing models according to one survey.

And another report indicates that only about 54% of models make the transition from pilot to full scale production.

And I think this is really kind of revealing a big need for kind of effective model testing solutions while industry giants, of course, we have like the Amazon, Google, Microsoft and those kind of companies, they all, you know, they offer similar services.

Other startups like Prolific, Robust Intelligence, Deep Cheat Checks, and Bobbidi are also kind of entering the arena with a bunch of different kind of innovative approaches that this is well.

However, Kalina claims to have an edge by offering customers full control over data types, evaluation

logics and some other testing components.

I think it's kind of robust focus on privacy eliminates the need for users to upload their

data or models.

And as Kalina only retains test results, which can be deleted upon request.

So looking ahead, I think while Kalina remains pretty discreet about its current customer

base, some people are saying that's because they don't have a good customer base.

Some people are saying, you know, it's just confidentiality reasons.

But I think the company is adopting a quote unquote selective approach focusing on partnerships

with mission critical enterprises.

Now really, like if I'm going to be honest, what I read between the lines on that is a

selective approach means they don't really have a ton of users.

That's not a bad thing.

That doesn't mean they're not a great company and they're not going to be, you know, successful.

But they probably don't have a massive user base.

Selective approach also could mean that their technology is maybe not robust enough to roll

out to a large larger audience, right?

Maybe they have had bigger companies and a big user base request to use them, but their

technology isn't strong enough to do that yet.

It's not bulletproof and they don't want to open themselves up.

So a selective approach means they can pick one or two clients that have very specific

needs kind of in the space.

They can make sure that they're fulfilling those needs and they're able to kind of like

test it, beta, get rid of bugs and, you know, remove, add the needed features and really

kind of test the waters.

I think it's actually a really good approach.

I have nothing against it, but I mean, I'm just, you know, trying to be transparent with

what I think that means.

In any case, the startup has also plans to introduce team bundles aimed at mid-sized

organizations and budding AI startups in the second quarter of next year.

And they said, quote, minimizing risks from an AI and machine learning system requires

rigorous testing before deployment, yet enterprises don't have strong tooling or processes around

model validation.

Co-Lenna focuses on comprehensive and thorough model evaluations.

We give machine learning managers, product managers and executives unparalleled visibility

into a model's test coverage and product-specific functional requirements, allowing them to

effectively influence product quality from the start.

So I think as AI really continues to kind of permeate every facet of our lives, tools

that can evaluate, validate and benchmark AI models not only become like really indispensable

but they also might kind of dictate the future credibility of the technology, especially specific

models.

So I think with this new funding, Co-Lenna seems well positioned to play a pivotal role

in this evolving kind of landscape and I'm really excited to follow them on their journey and see where they go.

If you are looking for an innovative and creative community of people using ChatGPT, you need to join our ChatGPT creators community.

I'll drop a link in the description to this podcast.

We'd love to see you there where we share tips and tricks of what is working in ChatGPT.

It's a lot easier than a podcast as you can see screenshots, you can share and comment on things that are currently working.

So if this sounds interesting to you, check out the link in the comment.

We'd love to have you in the community.

Thanks for joining me on the OpenAI podcast.

It would mean the world to me if you would rate this podcast wherever you listen to your podcasts and I'll see you tomorrow.