

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

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 11Labs, 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 waitlist 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.

Welcome to the AI Chat podcast, I'm your host, Jayden Schaefer.

Today on the podcast, we have the pleasure of being joined with Fedor Shtanov, who is a seasoned AI expert with over 18 years of experience in machine learning and product development.

Formerly a principal scientist at AWS, he now serves as the head of AI at Toloko, who is focusing on creating state-of-the-art responsible AI technologies with human oversight.

A thought leader in the field, Fedor has contributed to VentureV and is really passionate about bringing bridging kind of the gap between machine learning and human insight.

Welcome to the show today.

Thanks a lot, dear.

Thanks for having me.

Thank you.

Super excited to have you on the show.

I wanted to kick this thing off and kind of ask you a little bit about your journey.

If you could walk us through what that looked like, you were a principal scientist at AWS and now you're kind of the head of AI.

Did you always know that you were interested in AI and kind of this tech space growing up or is this something that you discovered as you were going through school?

Tell us a little bit about your background.

Well, actually, Gioski was specifically built at AI.

I was doing some applied mathematics and numerical methods and then I found a scientific supervisor

who was like, hey, there's this machine learning thing and it's pretty cool.

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

And I was back in 20 or something.

And everybody was using sports factory machines at the time.

And we found a really cool application for a sports vector machine to plasma physics.

Well, this is a really neat way of automating really complex discoveries, which is due and the journey started and I did my PhD in computer science and worked in a few companies. Should Amazon, Microsoft, AWS.

But what's also curious and I actually hope that machine learning folks can hear me because I think that's worthwhile.

At some point, I realized that everybody in the eyes about automation and everybody is like, okay, let's automate this, let's automate that.

But the real beneficiary out of all of this showed that at the end of the day, be humans and how do you make the AI really beneficial for humans?

How do you make sure that humans stay in the loop when the AI makes decisions, how do you make sure that the AI makes decisions which are useful?

And this is the area where that kind of shifted a few years ago, actually about five or six years ago.

And in AWS, I was working in the same area as well.

And so, again, from AWS to Telok to progress in that area of human loop and now I'm head over behind Chief Scientist.

That's super cool.

I'm wondering, could you tell us a little bit, it sounds like an incredible background and you've been working in this for a long time, so you've seen this whole thing evolve a lot.

Can you tell us a little bit about what got you interested in Telok moving, right?

You obviously probably had a great job over at AWS, I know a bunch of folks that work there and it's a fairly good place to be, whatnot.

What kind of got you interested in moving to this kind of new company and this new idea?

Yeah, yeah, actually happy to share this.

Imagine that you're doing artificial intelligence and you're thinking, well, how do I make it useful?

How do I make my models?

And then you very often before the age of the large language models, you often had to collect label data.

You had to say, you know, whether it's good, which is bad, and you had to go through this process of collecting data, which is getting seen back from humans.

And very often you see in the papers where it's hard, expensive, it's a manual process where you have to work with humans and not math and algorithms and for tech people sometimes and stuff.

And then you realize that there is a theory, which is crowdsourcing, which is basically like you have humans all around the world, who some of them have nothing to do with some of them, you know, really want to earn a little bit of money on the side, some of them want to earn a fair amount of money.

And then you think, how do I encode them in the process of making AI?

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

And so this is where Toloka, in my opinion, excels, is really incredible. It's better than many, many others, if not better than everyone is doing this job of involving humans around the world into the AI making process. And this becomes a much more technological problem than the problem of normal human collaboration or all human management. So it's a technological problem to, you know, to collecting, working with humans, a technological way of working with humans, a technological solution to that problem. And this is what is really cool, I think, in Toloka. It's a technological platform for dealing with humans and contributing to the AI. Super fascinating background. I love what you are working on. I'm wondering if you could share some success stories where human in the loop processes significantly improved AI performance. Yeah, let me recall a couple of stories from the background. One thing which is certainly worth noticing is that humans are always present in improving AI performance. You use a basically crucial degree. This is like how AI is built. For LLMs, you use humans, as I mentioned, to provide ground truth labels on which AI straightened and says, it's good, this is bad, humans work, those who fight it in the loop. Now with LLMs, you still have the same thing. You have this alignment process where you ask LLMs to produce results which sound better than humans, which actually reply to human queries. So without the human involvement, you very early get used to the layout. But if you think about the human in the loop, in the usual way, people understand it, where the models and machines at the same time work together, which is like a better benefit. The most popular example is probably the moderation example, the example where let's say there's a lot of user content like text to some forms or maybe some images are submitted to a website and then you need to filter out in real-life, in real-time, bad images, inappropriate images or bad content. So you usually build a model and then very often the model doesn't give you high enough ideas. So you trust the model when the model is confident, you don't trust the model when the model is not confident, you rely on humans to fill this gray area, like skis. This one is, now the more fun examples are, if you think about Amazon Go, if you're not familiar with it, if you haven't heard about it, it is a cashierless store. So you enter the store, you take what you want and you leave. The thing that works is that there's a lot of Canvas in the ceiling of that store and the truck, the person who enters, they see what they take. Right.

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

So that's the way it works, that there's machine learning which tracks you, which detects what you took and then basically sends you a receipt afterwards to a new store. So this happens very quickly after you get the store, you get the receipt, most of it. Sometimes if you toss your product to someone else or if you try to trip the system where you give it to someone else and leave it with that product or you try to put the product in one place and then you take it from the wrong place, it took, a few years ago, it took quite some time before you get the receipt.

And the reason for that is that the complex cases were monitored, were actually observed by the real humans and then they were checking.

Oh, okay.

And that's a more of a fun example.

Then there's a lot of fun examples.

When Yandex was building a personal assistance, it's something similar to Amazon Alexa or Google's personal assistant, Alice.

And before they even released, just when they started collecting data, basically the engineers were given this chatbot, like a preliminary version of a chatbot to ask this chatbot questions and then the chatbot would give them answers.

And the engineers were talking to that chatbot thinking that they're talking to some really better version of the model and then someone decided to find out, oh, where does this model hosted?

What's the computer?

And then they didn't find the computer when the model was hosted and they figured out that the whole person was actually building a way that the engineers asked the question.

It's real people, assessors who answered those questions, versus updated collection.

That's another fun story of how human instead of AI helps in getting used to things.

That's super cool.

Yeah, those are awesome stories.

Really interesting.

I had no idea.

So that's some cool insights.

I'm wondering, how does Toloka ensure the quality of human-labeled data across such diverse languages and countries and all that?

Yeah, sure.

That's a good question.

As I mentioned before, Toloka approaches this way of the problem of getting data or interacting with humans through a technological perspective, thinking about what kind of abstractions,

what kind of tools can help you get higher quality data.

And so there is actually quite a variety of popular cross-sourcing tools which help you get a higher quality data.

So first you give people entry tips.

Some tasks where you don't want people to waste their time, you actually give some sort of fairly simple museum tests for how well they could pass it.

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

Do they know English?

Do they know Spanish?

Do they know...

On a server then.

Then you want to actually get some higher proficiency, give people training, and then there is...

This is actually giving people training, it's not a manual process, you design training tasks, training tasks for people where you say, this is a task, please follow it, and then say, performance, and say, no, you're incorrect, we chose this answer, but you should have chosen that answer, this is why.

So people learn this way, and then they get through a second test or exam, which is of course higher quality or like insurance that you can perform, you know, the person can perform it at a level, and so then the person is allowed to perform real tasks.

However, there are some people, of course, who are trying to cheat through the system, and who are trying to get through training and exam with themselves, and then they try to automate or not really get cancer.

So to avoid this, we have other tools, which is we actually have some coupadis parts where basically somewhere in the middle of your answers, there's an answer with a known...

There's a question with a known answer, and if the person answers incorrectly too many times to those kinds of questions, they're not allowed to perform the tasking.

And so incorporating those honeypots into the process of labeling, into checking, regular checking of performance is a crucial part of high quality labeling.

Those honeypots are built by more trustworthy people, most trustworthy assessors who have been tested by themselves by even higher, more trustworthy experts, to basically have this pyramid of quality feedback providers.

And so the more expert cheese you require from a provider, usually the rare chance you have to find them.

And so as a result, they have a lot of people who have to be checked by trustworthy people and those have to be checked by experts.

So this is the way we approach our getting records.

Okay, very interesting.

That's a problem I had not really thought about the solution to solve.

It's fascinating how you guys are currently working on it.

With all of this that you're currently working on, what do you think are some of the ethical challenges that maybe yourself, but also the AI community needs to address urgently?

What's kind of on your radar right now?

So to me, there's a lot of discussions right now about, you know, Jerry Lee, about large language models, about putting some sort of, some sort of government restrictions on who can provide the models.

In my opinion, I can build those models.

In my opinion, what's quite problematic is that right now the reason that much of a variety of models, which are high performing, and which are, which can be easily representative of the various groups of people, current high performing models are basically based on what

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

frankly buys there and the data for which what is good data, what is bad data is determined by a very small group of people who quite often do not represent some other groups which are interesting, some other groups which are important.

And in my opinion, what's worth doing is actually being well open about how you construct a task to large language models of these foundation models, sharing instructions or sharing the actual data for fine tuning or aligning those models in a way that other interested members can build the models which are less wise to that specific small group.

So that's quite important because otherwise it's influencing the whole society.

You have seen the amount of people who are starting to be using chat GPTs, quite a lot of people, the influence also.

What kind of answers the models give is quite significant.

Yeah.

Yeah.

No, I think that's super important.

It's something I definitely think a lot about because of course right now we have, like you mentioned these large language models, we have OpenAI, we have Google and Meta and really all of these companies coming up with these, it's kind of a small bubble around San Francisco really.

And it's like we all know that San Francisco has a lot of its own ideas.

Some people love some of them, some people don't love all of them, but there's definitely different ideas all around the world that different people subscribe to or don't subscribe to.

And so yeah, having one small group of people create these essentially large models that like for Googles, you know, they're integrated in their search, this is going to be seen by billions of people around the world.

Where do you see this going in the future in regards to that?

Do you think that there's a place where someone like OpenAI or Google or Meta could create a model everyone would want?

Or do you think, you know, at some point people are just going to say, look, I don't align with these large models on XYZ topics.

I'm going to find maybe like there's, you know, thousands of models and this is the kind of the one that more aligns with my beliefs.

Do you see the industry going in that direction?

So I, that's a great, I wouldn't have the image of question because people are betting either the model of open source or that the open source is going to grow and which is going to have those models available or betting that pay and GenAI is the next cloud and that is, it's a great question.

I tend to be on a camp which thinks that all the time, if you look at the history of machine learning and things like building models between cheaper and over time, all it always becomes cheaper.

It always becomes more accessible to, you know, quite large variety of different groups. And the information of how to build those models also leaks because people need one company and get another.

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

So I really hope that the future is going to be that the technology is going to spread and more people are going to be able to build those kind of models and tune them to whatever needs they see, whatever underrepresented groups they see in the society so that you have a diversity and not basically a single or a few heads of pity about the topic.

So this is, this is my, this is my hope.

I do think that the openness helps here, but I do think also that it's worth being careful about opening everything is sort of putting some sort of guards or government laws against it is used or it is possible to waste this models to be used for some legal purposes.

So it's certainly worth doing that.

But I did hear in the, you know, in the industry someone quite radical in my being propositions of say allowing the models to be only created by very, very specific list of say companies or institutions.

And that to me seems quite a, quite a risk empathetic.

Yeah.

Yeah.

I 100% agree.

It's interesting.

You know, open AI, of course, I'm all and famously went to Congress in the United States and said, we need to regulate AI.

I'll help you write the regulation.

We need to decide who gets to, you know, get a license to train the models.

And of course, a lot of people raise red flags about that when, you know, the biggest company with \$10 billion says that they want to be the one to help give out the certificates of who can and can't do it.

So I think they'll be pushed back.

Of course, there's the open source community.

I think a lot of the stuff meta has done has honestly, it's kind of interesting.

I'm not usually like a huge fan of meta and Facebook and whatnot.

But some of the open source stuff that they've done with AI, I think is really interesting and I got it.

I just have to, you know, give them credit or credit to do on that.

And even a lot of the infrastructure they built.

So very cool.

See, like kind of what we're seeing in this space, I think, inevitably, whether there's regulation or not, like if there's too much regulation, people are just going to get like boot, you know, bootlegs, open source models, you're going to have them on a thumb drive and it's going to be your, you know, open source model that doesn't have the open AI guardrails or, or whatever, you know, I don't know, it'll be interesting to see what goes in the case.

I guess based on your really extensive experience you've had in this space, something I would love to ask you about as we kind of wrap up the show is, you know, what advice would you give to companies looking to implement AI solutions responsibly today?

I'm glad you asked about the responsibly part because that's I think it's quite important

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

nowadays with the models don't only give you the final answers.

This is good.

This is bad.

Your mistake is costly for the binary models, but it's not nearly as costly.

It can be not nearly as costly as the mistake for generated ones, which significantly offensive.

I mean, you know, can really provide some really unpleasant or illegal justifications or descriptions and because of this larger impact of the just specifically the generating models, because of how much risk they have.

I'd say that it's really important you're developing some sort of a product based on generating AI, not to save on evaluation, not to, you know, evaluate your users, tell me what's good and what's bad, really investing in evaluation.

Really think about in advance, how do they judge whether the model is not providing risk answers, not providing some unbiased opinion, which reinforces some sort of societal problems, but really, you know, provides great experience where getting real tools instead of making up facts or collocinating what they call.

And so it should be done in my opinion, really early in the process of model development. So I'd say that this is, she asked me about single advice, I'd say that if I usually hear you worth it, and it proves that you take things responsibly, you actually think about what you built and you're sure about it and you can demonstrate.

I love that.

I think that's some really solid advice.

Flore, thank you so much for coming on the podcast today.

I really appreciated a lot of your advice and insights.

It's really a refreshing because I get a lot of different, you know, perspectives on this show, and I really think you're on the right track with, you know, what you guys are working on, what you're doing and kind of your philosophy behind it. So that's always, that always makes me happy to hear.

If people are interested in getting in contact with you or Toloka and finding out more about what you guys are building and working on, what's a good way for them to do that?

Yeah, they can connect to LinkedIn or I think that's the best way, just finding on LinkedIn to connect, right?

OK, and thanks for having me.

It was a great conversation.

It was interesting and on an active.

Yeah. Yeah, it was super enlightening.

So really appreciate you coming on to the listeners.

Thanks so much for tuning in to the AI Chat podcast.

Make sure to rate us wherever you get your podcasts and have a wonderful rest of your day.

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

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / ChatGPT's App Hits \$4.5M Monthly Revenue, But Growth Slows

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 chat GPT. 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 Open AI 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.