

## [Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / EPIK's AI Yearbook App Soars to #1 on App Store!

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 right now, of course, we have kind of this backdrop of a super strained AI chip supply chain, right?

This is something no one can get enough chips.

There's huge wait lists are very hard to get.

And I think amid all of this open AI, which is of course the number one AI startup right now is reportedly contemplating an entry into the AI chip space.

So this move comes as the demand for powerful AI processors to train their ever advancing models, right?

We've got chatGPT.

GPT-4 is now the big thing and GPT-5 is rumored to be coming out soon.

Well, I don't know what the rumors are, but I think it's going to come out in December and you can just, you know, I mean, that's my prediction based off of a bunch of stuff I've seen them trademarking some things around the space and some people saying that they're already in the training process could be completely wrong.

But I think they might just hit it one year later in December, GPT-5, but in any case, they're also just they've just launched Dolly 3, which is, you know, their upgraded image generator, which I think is a really good move.

In any case, of course, all of these things are sucking a ton of compute.

So OpenAI currently depends on GPU based hardware, like a lot of its different, you know, peer companies.

And I think right now they may be looking to change that because GPUs have been essentially the backbone for AI development due to their proficiency in handling parallel computations, which is really a necessity for training today's kind of top tier AI.

## [Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / EPIK's AI Yearbook App Soars to #1 on App Store!

But the rise of generative AI, of course, has really, really helped out companies like NVIDIA who are making money off of the chips.

So however, I think it has kind of pushed this whole GPU supply chain to its limits.

Microsoft recently cautioned about potential service disruptions due to like issues with server hardware because of AI.

And also NVIDIA's high performance AI chips are reportedly, you know, sold out until 2024.

So really, it's it's hard to get your hands on those things.

And I think a lot of these AI companies are kind of feeling the strain because it's not just OpenAI that needs its own chips.

There's all these other companies that are doing and it probably feels like they're getting in their own way.

And it's like kind of a bottleneck to their own company.

So I'm sure OpenAI at this point is like, you know, doesn't want to put their company in the hands of another enterprise.

So it looks like they're kind of eyeing this space.

So to give a little bit of context and a maybe perspective on kind of the magnitude of demand, a study by Bernstein analyst, Stacey Raskin highlighted that if chat GPT queries scale to even a tenth of Google searches volume, an initial invest investment of roughly \$48 billion in GPUs would be necessary.

This is insane.

And I'm going to say that again, because I do not think people fully understand this.

We talk about like, oh, chat GPT is awesome, yada, yada, whatever.

And Google like should be worried it's going to get beaten and Google is kind of doing this stuff with Bard on the side.

But like, if Google searches really were replaced by chat GPT, which to be honest, for a large part, I'm doing a lot of, I'm doing a lot of stuff on chat GPT that I would have used Google search for in the past.

But if that was to replace it, right, 10% of Google searches would require almost \$50 billion in GPUs to be purchased in order for OpenAI to facilitate that.

So that's absolutely insane.

If they were to do 100% of Google searches, we're talking about somewhere like \$500 billion.

This is insane numbers.

If Google was just a switch to completely be chat GPT instead of like a Google search, we're talking \$500 billion in GPUs that would be needed.

So I think another really interesting number is that around \$16 billion worth of chips would be required annually for sustained operations, right?

So even if they made like, if they say chat GPT is going to take over 10% of Google searches, \$50 billion up front on GPUs, and then every year \$16 billion.

This is insane money.

And I think chat GPT and OpenAI really know that if they want to be able to scale, this is a bottleneck for them, right?

Even if they were the most popular thing in the world, this is a bottleneck.

And I think they're looking at, you know, how can we build things?

## [Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / EPIK's AI Yearbook App Soars to #1 on App Store!

How can we build them better, cheaper so that we're not kind of stuck with this bill? Even if they could get at half price by making it themselves, it would make a really big difference.

So OpenAI is by no means kind of pioneering the custom AI chip territory. Apple's TPU tensor processing unit powers massive generative AI systems like Palm 2 and Imogen.

Amazon provides its AWS customers with proprietary chips designed for both training. I think it's called Trainium and Inference, which is Inferentia.

So meanwhile, there's a whole bunch of rumors around Microsoft collaborating with AMD on an AI chip called Athena, which OpenAI is like reportedly testing right now.

So I think with a venture capital injection of over \$11 billion, that's what OpenAI has at the moment.

And you know, I think they're nearing a \$1 billion in yearly revenue, like that's kind of where they're at right now.

They said, I think Sam Altman earlier this year said they forecast are going to hit about a billion dollars in revenue this year.

I think OpenAI's financing position appears like to be pretty solid.

It's fairly robust, but recent, of course, murmurs from the Wall Street Journal suggest that the possibility of a share sale is going to catapult OpenAI's secondary market valuation to around \$90 billion.

They're kind of looking at doing that right now.

So that would be very, very incredible.

However, I think the road to AI chip development isn't without its bumps.

Just last year, AI chipmaker GraphCore saw its valuation plummet by around a billion dollars post a Microsoft deal fallout, Microsoft never ended up going through with the deal and it really kind of hammered their company.

So I think that was kind of leading to an announcement of job cuts owing to challenging, you know, they said this like challenging economic situation stuff.

So I think the past month also saw GraphCore grappling with declining revenue and mounted losses.

Intel's AI chip subsidiary, which is called Habana Labs, had to let go of an estimated 10% of its employees.

Even tech giant Meta faced a bunch of turbulence with its custom AI chip endeavors, eventually axing some of its experimental hardware that they were kind of working on before.

So it's not like it's all roses in this space and it's not like it's very easy, right?

Of course, we have NVIDIA, which is absolutely crushing it, but there's a lot of people that are kind of hurting.

Maybe that's because an NVIDIA was significantly better.

There's a lot of people speculating on why that is.

But I think, well, OpenAI's potential move into the AI chip domain is a significant development.

The journey to bring a custom chip to market is going to be really lengthy and might drain hundreds of millions annually.

The critical question I think that remains is whether OpenAI's stakeholders, which of

## **[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / EPIK's AI Yearbook App Soars to #1 on App Store!**

course is Microsoft that owns half the company, are ready to kind of place their bets on such an ambitious and uncertain venture, especially when Microsoft's already, you know, inking deals with AMD and others.

I think only time is going to reveal if OpenAI's potential chip endeavors are actually going to be successful or not.

And the one thing I would just stress, finally, is that these these deals really do take a long time.

If they were to design their own chip, if they were to try to launch this, building out the facilities to fabricate it, building out, you know, the essentially all the manufacturing like this stuff takes years.

We'd be talking two, three, like two years would be fast, right?

We're talking probably four years before they could scale this thing out.

And so really at the moment, I think it's this is not a solution that really pulls them out of this problem.

But you know, maybe it's future revenue, maybe it helps them in the future.

But at the moment, I think everyone really is kind of stuck with NVIDIA and a couple of the other players and essentially just trying to make do with what they have currently on the market.

If you are looking for an innovative and creative community of people using chat GPT, you need to join our chat GPT 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 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 or wherever you listen to your podcasts and I'll see you tomorrow.