

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / OpenAI's ChatGPT: Now Seeing, Hearing, and Speaking!

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 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 the big news here, there was recently a demo released, a lot of people are talking about it and in the demo, essentially someone takes a picture of their bicycle, uploads it to chatGPT on their phone on the mobile app.

I believe this is iOS, I tried to download this on Android, I'm not seeing it.

So yeah, anyways, they took a picture of their bicycle on chatGPT, they were able to upload it and as soon as it was uploaded, they were then able to add the message saying, you know, help me lower my bike seat, right?

And so immediately after sending this, chatGPT is able to come up with a response that says, you know, to lower your bike seat, locate the release lever, you know, if it's a quick release lever, do this, tighten the quick release.

Anyways, it walks through the steps of doing it, it says, if you have further tools, show me and I'll guide you through.

So then the person goes, they take a picture of where the kind of release thing on the seat is, they circle it on their phone.

So they're actually like drawing attention to the specific part they're talking about and they say, is this the lever, chatGPT says, no, that's not the lever, it's a bolt, you'll need an Allen wrench to loosen it.

And then it, you know, asks about what tools they have, so they go and take a picture of tool set they have.

And they also take a picture of, I believe, the users or the owner's manual for the bicycle and they say, you know, here's the manual and here's my toolbox, do I have the right tools?

And this is like so fascinating because it's using this computer vision at the same time, it says yes, you have the right tools on the left section of your toolbox, there's a set

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / OpenAI's ChatGPT: Now Seeing, Hearing, and Speaking!

labeled the wall within the set, find the nine millimeter Allen head.

So it's literally like looking at the picture and it's like talking you through, like on the right side of your toolbox on that little thing with this label on it, like it's telling you what specifically to go and find, go and do, really, really cool feature here.

So of course, this is the ability that chatGPT now has to see, this is kind of something that was rumored and talked about back when GPT-4 launched, but it never actually happened at the beginning of the year, until now it looks like they're finally rolling this in.

It's kind of interesting, I recently saw a tweet on Twitter where someone was saying they're like, you know, OpenAI is a big conference planned for November, but what the heck is there left for them to announce, like AGI, because after this announcement, like honestly this is some massive capabilities, right?

So now you can take pictures, upload it to chatGPT and ask it about the picture, so it's literally able to see.

In addition, like I mentioned in the intro, there's the huge news of Spotify being able to do voices, or they're partnering with OpenAI to do voices, so they're going to translate podcast voices into different languages, so essentially clone a voice and translate it.

I've also seen demos with chatGPT where people were able to, you're able essentially to talk to it, so send an audio file where you are talking, you say, hey chatGPT, tell my kids a bedtime story about a hedgehog that lives on the top of a mountain, and chatGPT was able to say, sure, here's a bedtime story and talk back to you, so chatGPT can talk back to you, you can talk to chatGPT, you can take pictures, send it to chatGPT, ask it for context or information about the pictures that you're seeing, it can see, it can hear, and most recently, of course, with the huge announcement that they're going to start integrating the brand new version of Dolly 3 in there, so it's also able to train your really interesting pictures.

I saw some interesting demos where essentially they were talking to chatGPT and saying like, hey, you know, tell my kids a bedtime story about a hedgehog, it tells them the story about the hedgehog, right, using the voice feature, but then they're like, okay cool, my kids really want to know what the hedgehog looks like, can you show us a picture of what this hedgehog would look like if it was on top of a mountain in a forest, and so it generates the picture, and it's like, okay, cool, my kids are about to go to bed, can you generate a picture of this hedgehog as it's going to bed, it does that, same hedgehog, and this is another thing that I think is really interesting with something like Mid Journey, you can try to use the same prompt to get a very similar type of character, so for example, I've spoke with a number of people that have made like children's illustrated books, and they have like keywords they put in there to try to get the same characters in the books, that's one of the hard things is to make it really cohesive, but they say things like, you know, like illustrated in like X, Y, and Z style, and they know that if they say like girl with black hair in this style, it tends to come up with a very similar looking girl, and so they were able to make a bunch of kids books doing this, but they're not like perfect, and so they're not exactly the same person, it was like okay, right, but I think what's really incredible here that ChatGPT is going to have the capability of doing is because you can upload pictures and ask it questions about that, you can say, you know, generate

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / OpenAI's ChatGPT: Now Seeing, Hearing, and Speaking!

me the picture of this hedgehog, or this character in a book, or this person, or this thing, it generates it, and you say okay, now generate that same thing, but in this scene, in this place doing this thing, or that, and all of a sudden ChatGPT because of the way it's integrated with natural language processing, and it has this computer vision element where you're able to upload and talk, it's able to do some really powerful effects where it's like essentially manipulating the same characters, the same images to give you, you know, different variations of that, which are going to be really powerful. So many incredible use cases are coming out with this, it's all over Twitter, people are talking about some really impressive use cases, but honestly, you know, recently we even had Lex Friedman talking about, he's one of the first people that are going to be on this whole spotify thing, and he recently said, he posted a clip of literally him speaking in English on his podcast, and it translated to Spanish, he said, this is me speaking Spanish thanks to amazing work by Spotify, AI, engineers. The translation and voice cloning are done by AI languages, can create barriers of understanding, and thus fuel division. I can't wait for AI to break down this barrier and reveal our common humanity. Great job, Spotify, I'm excited for what the future holds, check out the full episode of me speaking Spanish. So, you know, I mean, he's trying to promote his podcast, which by all means, that's awesome, but you know, just then on LinkedIn had almost 10,000 people that have liked it in the last 21 hours. So really, really powerful technology, I think this kind of just goes to show like this is something people definitely want. I, even for myself, I've been invited on a number of podcasts where they're like, hey, I'd love to have you on this podcast, we talk about AI, but it's in like, you know, we're based in Sweden, and like that's pretty much our audience. And so like for me, it's just like not super viable. Obviously, I don't speak anything other than English and French, so I'm kind of limited to those two lanes. But I think this is incredible, being able to translate into all these different languages and clone your own voice. Absolutely incredible. I see, you know, if this is really what OpenAI is building in, you can see all of the different incredible use cases that are going to happen, you're essentially going to be able to go travel to a new place, and you know, talk into your phone, and it's going to in real time translate for people. So that or even maybe you wear a pair of AirPods with your phone that is listening. So when people talk to you in another language, it immediately translates into English, you speak into your phone, and then it's going to in your same voice, speak to them in French or Spanish or German or whatever, wherever you're visiting, and give them that, you know, you're going to be able to communicate so much easier, there's going to be some incredible applications built on top of this technology. And of course, on top of the technology that we're able to see things all around us. So, you know, imagine traveling, you're in another country, you take a picture of a street, and you're like, hey, I'm looking for like, you don't know what any of the shops are. But it's like, based on all the street signs, which one of these is like, which one of these is like a shoe store or something, right? I don't know, it sounds funny, but there's going to be so many incredible use cases. I think this is going to just further integrate AI into things. And honestly, I think right now, this is OpenAI is kind of secret to getting ahead, because they're building APIs with all of this stuff. And so if you think about it, you know, right now we have like Anthropic that just raised \$4 billion, or has the ability to raise up to \$4 billion from Amazon and AWS, you know, they're

[Transcript] AI Hustle: News on Open AI, ChatGPT, Midjourney, NVIDIA, Anthropic, Open Source LLMs / OpenAI's ChatGPT: Now Seeing, Hearing, and Speaking!

still just trying to compete on chat GPT alone, and just that tool. But as chat GPT now is starting to integrate this image generation and the speech and this audio and all sorts of other things. And of course, I, you know, I find it hard to believe that video is very far away. I think, you know, a lot of their competitors are going to have to fight really hard to still be relevant, because I think once people get all of these new features from chat GPT, it's going to be really hard to convince yourself to go and pay for something like Anthropic, if it doesn't have those abilities to have, you know, hearing, seeing, speaking, and everything else integrated, it's going to become incredibly convenient. And I think this may kind of push it to mainstream beyond just people using chat GPT to, you know, generate articles or responses and texts for them. But all of a sudden, it may be just an app everyone has to have, everyone has to have premium and you use it all the time. It's like your travel companion, you can't live without it. I see it going more in that direction with some of these new features. So very exciting, definitely something will continue to follow and a lot of amazing advancements to come. 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 wherever you listen to your podcasts and I'll see you tomorrow.