

URL: <https://ecorner.stanford.edu/videos/inside-openai-entire-talk/>

Ilya Sutskever is the co-founder and chief scientist of OpenAI, which aims to build artificial general intelligence that benefits all of humanity. He leads research at OpenAI and is one of the architects behind the GPT models. In this conversation with Stanford adjunct lecturer Ravi Belani, Sutskever explains his approach to making complex decisions at OpenAI and for AI companies in general, and makes predictions about the future of deep learning.



Transcript

- Who you are (electronic music) defines how you build. - Welcome YouTube and Stanford communities 00:00:11,910 to the Entrepreneurial Thought Leaders Seminar, brought to you by STVP, the Entrepreneurship Center in the School of Engineering at Stanford, and BASES the Business Association of Stanford Entrepreneurial Students. Today we are so honored to have Ilya Sutskever here at ETL. Ilya is the co-founder and chief scientist of OpenAI, which aims to build artificial general intelligence for the benefit of all humanity. Elon Musk and others have cited that Ilya is the foundational mind behind the large language model, generative pre-trained transformer 3, or GPT-3, and its public facing product ChatGPT. Few product releases have created as much excitement, intrigue, and fear as the release of chat GPT in November of 2022. Ilya was, Ilya is another example of how the US, and the world, has been the beneficiary of amazing talent from Israel and Russia. Ilya was born in Russia, and then when he was five, he moved to Israel where he grew up. And he spent the first half of undergrad even in Israel. And then he transferred and went to the University of Toronto to complete his bachelor's degree in mathematics.

He went on to get a master's and PhD in computer science from the University of Toronto, and then came over here to The Farm, and did a short stint with Andrew Ng before returning back to Toronto to work under his advisor, Geoffrey Hinton's research company, DNN Research. Google then acquired DNN Research shortly thereafter in 2013, and Ilya became a research scientist at it as part of Google Brain. And in 2015, he left Google to become a director of the then newly formed OpenAI. It's hard to overestimate the impact that ChatGPT has had on the world since its release in November of last year. And while it feels like chatGPT came out of nowhere to turn the world on its head, the truth is there's a deep history of innovation that has led to that moment. And as profound as ChatGPT is, Ilya is no stranger in ushering in discontinuous leaps of innovation in AI. Jeff Hinton has said that Ilya was the main impetus for AlexNet, which was the convolutional neural network in 2012 that is attributed to setting off the deep learning revolution that has led to the moment that we are now in. And of course, it was seven years since the founding of Open AI that ChatGPT was finally unleashed to the world. Ilya was elected a fellow of the Royal Society in 2022. He's been named to the MIT Tech Review 35 under 35 list in 2015.

He's received the University of Toronto's Innovator of the Year Award in 2014, and the Google Graduate Fellowship from 2010 to 2012. So with that, everybody, please give a virtual warm round of applause and welcome for Ilya to the Entrepreneurial Thought Leader seminar. So Ilya, imagine lots of applause, and you're always invited back onto The Farm

physically whenever you are able. So Ilya, there's so much to discuss, and I know we're gonna have so little time, we have quite a broad range of fluency around the audience in terms of ChatGPT and large language models. I wanted to start off with just a quick question on the technology, which is, just, the key technology underlying OpenAI and generative AI more broadly, is large language models. Can you describe the technology in simple terms? And now that you're at the forefront of the tech, can you share what has surprised you the most about what the tech can do that you didn't anticipate? - Yeah, I can explain 00:03:47,163 what this technology is and why it works. I think the explanation for why it works is both simple and extremely beautiful, and it works for the following reason. So you know how the human brain is our best example of intelligence in the world. And we know that the human brain is made out of a large number of neurons, a very, very large number of neurons. Neuroscientists have studied neurons for many decades to try to understand how they work precisely.

And while the operation of our biological neurons are still mysterious, there's been a pretty bold conjecture made by the earliest deep learning researchers in the forties. The idea that an artificial neuron, the ones that we have in our artificial neural networks, kind of, sort of similar to a biological neuron if you squint. So that's, there's an assumption there. And we can just run with this assumption. Now, one of the nice things about these artificial neurons is that you can, they're much simpler, and you can study them mathematically. And a very important breakthrough that's- that was done by the very, very early deep learning pioneers before it was known as deep learning, was the discovery of the back propagation algorithm, which is a mathematical equation for how these artificial neural networks should learn. It provides us with a way of taking a large computer and implementing this neural network in code. And then there would be, there is an equation that we can code up that tells us how this neural network should adapt its connections to learn from experience. Now, a lot of additional further progress had to do with understanding just how good and how capable this learning procedure is, and what are the exact conditions under which this learning procedure works well. It's, although this is, although we do with computers, it was a little bit of an experimental science, a little bit like biology.

We have something that is, you know, like a biological experiment a little bit. And so then a lot of the progress with deep learning basically boils down to this. We can build these neural networks in our large computers, and we can train them on some data. We can train those large neural networks to do whatever it is that the data asks them to do. Now the idea of a large language model is that if you have a very large neural network, perhaps one that's now not that far from, like these neural networks are pretty large, and we train them on the task to guess the next word from a bunch of previous words in text. So this is the idea of a large language model. You train a big neural network to guess the next word from a previous, from the previous words in text. And you want the neural network to guess the next word as accurately as possible. Now the thing that happens here is we need to come back to our original assumption that maybe biological neurons aren't that different from artificial neurons. And so if you have a large neural network like this that guesses the next word really well, maybe it'll be not that different from what people do when they speak, and that's what you get.

So now when you talk to a neural network like this, it's because it has such a great, such an excellent sense of what comes next, what word comes next, it can narrow down, it can't see the future, but it can narrow down the possibilities correctly from its understanding. Being able to guess what comes next very, very accurately requires prediction, which is the way you operationalize understanding. What does it mean for a neural network to understand? It's hard to come up with a clean answer, but it is very easy to measure and optimize the network's prediction error of the next word. So we say we want understanding, but we can optimize prediction, and that's what we do, and that's how you get this current large language models. These are neural networks which are large, they're trained with a back propagation algorithm, which is very capable. And if you allow yourself to imagine that an artificial neuron is not that different from a biological neuron, then yeah, like our brains are doing are, are capable of doing a pretty good job at guessing the next word. If you pay, if you pay very close attention, so- - So if I, I love that, 00:08:50,700 and I just wanna make this more concrete. So just to push that analogy further between the biological brain and these neural analog physical networks, digital networks, if the human, if we consider, you know, before it was considered untenable for these machines to learn, now it's a given that they can learn or do this, do predictive outcomes of what's gonna come next. If a human is at 1 x learning, and you have the visibility into the most recent ChatGPT models, what would you put the most recent ChatGPT model as a ratio of where the humans are at? So if humans are at 1 x, what's ChatGPT at? - You know, it's a bit hard 00:09:34,140 to make direct comparisons between our artificial neural networks and people because at present, people are able to learn more from a lot less data. This is why these neural networks like Chat GPT are trained on so much data, to compensate for their initial slow learning ability.

You know, as we train these neural networks and you make them better, faster learning abilities start to emerge. But overall, overall it is the case that we are quite different. The way people learn is quite different from the way these neural networks learn. Like one example might be, you know, these neural networks, they are, you know, solidly good at math or problem meaning, but like the amount of math books they needed to get, let's say, good at something like calculus is very high. Whereas a person would need a fairly, you know, two textbooks and maybe 200 exercise and you're pretty much good to go. So there is- - But just to get an order of magnitude sense, 00:10:51,090 if you relax the data constraint, so if you let the machine consume as much data as it needs, do you think it's operating at, like, one-tenth of a human right now or. - You know, it's quite hard to answer that question still. 00:11:03,840 And let me tell you why I hesitate to give, like I think that any figure like this will be misleading, and I wanna explain why. Like, because right now any such neural network is obviously very superhuman when it comes to the breadths of its knowledge, and to the very large number of skills that these neural networks have. For example, they're very good at poetry, and they're very, they know, like they can talk eloquently about any topic pretty much, and they can talk about historical events and lots of things like this.

On the other hand, people can go deep, and they do go deep. So you may have an expert, like someone who understands something very deeply, despite having read only a small amount of documents, let's say, on the topic. So because of this difference, I really hesitate to answer the question in terms of, oh yeah, it's like some, some number between zero and one. - But do you think there is a singularity point, 00:12:06,990 where the machines will surpass the humans in terms of the pace of learning and adaption? - Yes. 00:12:15,840 - I don't know when it'll occur. 00:12:17,598 I think some additional advances will need to, will happen. But you know, I absolutely would not bet against this point occurring at some at some point. - Can you give me a range? 00:12:29,193 Is it at some point next month? Is it next year? - You know, 00:12:35,700 I think it's like, the uncertainty on these things is quite high because these advances, I can imagine it can taking quite a while. I can imagine it can taking a disappointing a long time. I can also imagine it's taking, you know, some number of years, but- - Okay, that's fine.

00:13:04,350 - And I know there's lots of push forward, 00:13:06,330 so I'm gonna ask this, like, one more question and then move on to some of the other issues, but, I know, I read that when you were a child you were disturbed by the notion of consciousness, and I wasn't sure what that word meant, disturbed. But I'm curious, do you view consciousness, or sentience, or self-awareness as an extenuation of learning? Do you think that that is something that also is an inevitability that will happen or not? - Yeah, I mean, on the consciousness questions, 00:13:35,550 like, yeah, I was, as a child I would, like, you know, look into my, at my hand and I would be, like, how can it be that this is my hand, I get to see like, I, something of this nature. I don't know how to explain it much better. So that's been something I was curious about. You know, it's, it's tricky with consciousness, because how do you define it? It's something that eluded definition for a long time, and how can you test it in a system? Maybe there is a system which acts perfectly right, but per perfectly the way you'd expect a conscious system would act, yet maybe it won't be conscious for some reason. I do think there is a si very simple way to- there is an experiment which we could run on an AI system, which we can't run on, which we can't run just yet. But maybe in like the future point, when the AI learns very, very quickly from less data, we could do the following experiment. Very carefully, we'd very carefully curate the data such that we never ever mention anything about consciousness, we would only say, you know, here is a ball and here's a castle and here is, like, a little toy. Like you would imagine you'd have data of this sort, and it would be very controlled. Maybe we'd have some number of years worth of this kind of training data.

Maybe it would be, maybe such an AI system would be interacting with a lot of different teachers, learning from them, but they're all very carefully, you never ever mention consciousness. You don't talk about, people don't talk about anything except for the most surface level notions of their experience. And then at some point you sit down this say AI and you say, okay, I want to tell you about consciousness. It's the thing that's a little bit not well understood. People disagree about it, but that's how they describe it. And imagine if the AI then goes and says, "Oh my god, I've been feeling the same thing, but I didn't know how to articulate it." That would be, okay, that would be, definitely, something to think about. It's like, if the AI was just trained on very mundane data around objects, and going from place to place or maybe, you know, something like this from a very narrow set of concepts, we would never ever mention that. And if it could somehow eloquently and correctly talk about it in a way that we would recognize, that would be convincing. - And do you think of it as a some, 00:16:07,980 as consciousness as something of degree, or is it something more binary? - I think it's something that's more, a matter of degree. 00:16:23,403 I think that like, you know, let's say if a person is very tired, extremely tired and maybe drunk, then perhaps if that's, when someone is in that state, and maybe their consciousness is already reduced to some degree, I can imagine that animals have a more reduced form of consciousness.

If you imagine going from, you know, large primates, maybe dogs, cats, and then eventually you get mice, you might get an insect like, feels like, I would say it's pretty continuous, yeah. - Okay, I wanna move on even though I could, 00:16:58,950 I would love to keep asking you more questions along the lines of the technology, but I wanna move on to talking about the mission of OpenAI and how you perceive, or any issues around ethics, and your role as chief science officer, how ethics informs, if at all, how you think about your role. And so lemme just lay a couple foundation points out, and then have you speak. As you know, OpenAI's mission is to ensure that artificial general intelligence benefits all of humanity, and it started off as a nonprofit and open sourced, and it is now a for-profit and closed source, and with a close relationship with Microsoft. And Elon Musk, who I believe recruited you to originally join OpenAI and gave \$100,000,000 when it was a nonprofit, has says that the original vision was to create a counterweight to Google, and the corporate world. And he didn't want to have a world in which AI which, is, has, which he perceives, and others, can have an existential threat to humanity to be solely in the holds of a corporate, of a for-profit. And now, OpenAI is neither open nor exclusively a nonprofit. It's also a for-profit with close ties to Microsoft. And it looks like the world may be headed towards a private duopoly between Microsoft and Google. Can you shed light on the calculus to shift from a for-profit to a nonprofit? And did you weigh in the ethics of that decision, and do ethics play a role in how you conceive of your role as the chief science officer? Or do you view it more as something that somebody else should handle, and you are mainly just tasked with pushing the technology forward? - Yeah, so this question is many parts.

00:18:44,850 Let me, lemme think about the best way to approach it. So there are several parts. There is the question around open source versus closed source. There is a question around nonprofit versus for-profit, and the connection with Microsoft, and how to see that in the context of Elon Musk's recent comments. And then the question about how I see my role in this, maybe I'll start with that because I think that's easier. - Okay. 00:19:30,273 the way I see my role, I feel a lot, I feel direct responsibility for what OpenAI does. Even though I, my role is primarily around advancing the science. It is still the case I'm one of the founders of the company, and ultimately I care a lot about OpenAI's overall impact. Now I want to go, so

with this context, I want to go and talk about the open source versus closed source, and the non-profit versus for profit.

And I wanna start the open source versus closed source. Because I think that, you know, the challenge with AI, is that AI is so all encompassing, and it comes with many different challenges. It comes with many different dangers which come into conflict with each other. And I think the open source versus closed source is a great example of that. Why is it desirable, or let me put it this way, what are some reasons for which it is desirable to open source AI? The answer there would be to prevent concentration of power in the hands of those who are building the AI. So if you are in a world where, let's say, there is only a small number of companies, you might let control this very powerful technology, you might say this is an undesirable world, and that AI should be open, and that anyone could use the AI, this is the argument for open source. But this argument, you know, of course, you know, to state the obvious, there are near term commercial incentives against open source. But there is another, longer term argument against open sourcing as well, which is if we believe, if one believes that eventually AI is going to be unbelievably powerful, if we get to a point where your AI is so powerful, where you can just tell it, "Hey, can you autonomously create a, like, I don't know, a biological research lab autonomously, do all the paperwork, run the space, hire the technicians, aggregate the experiments, do all this autonomously." Like, that starts to get incredible. That starts to get like mind-bendingly powerful. Should this be open sourced also? So my position on the open source question is that I think that, I think that there is maybe a level of capability, you can think about these neural networks in terms of capability, how capable they are, how smart they are, how much, how many, how much can they do? When the capability is on the lower end, I think open sourcing is a great thing, but at some point, and you know, there can be debate about where the point is, but I would say that at some point the capability will become so vast that it'll be obviously irresponsible to open source models.

- And was that the driver behind closed sourcing it, 00:23:03,330 or was it driven by a devil's compact or business necessity to get cash in from Microsoft or others to support the viability of the business? Was the decision making to close it down actually driven by that line of reasoning? Or was it driven by more financial? - So the way I'd articulate 00:23:23,760 it, you know, my view is that the current level of capability is still not that high, where it'll be the safety consideration it will drive the closed sourcing the model, the, this kind of research. So in other words, I claim that it goes in phases. Right now, it is indeed the competitive phase, but I claim that as the capabilities of these models keep increasing, there will come a day where it will be the safety consideration that will be the obvious and immediate driver to not open source these models. So this is the open source versus closed source, but your question had another part, which is non-profit versus for-profit. And we can talk about that also. You know, indeed it would be preferable in a certain meaningful sense if OpenAI could just be a for-a non-profit from now until the mission of OpenAI is complete. However, one of the things that's worth pointing out is the very significant cost of these data centers. I'm sure you're reading about various AI startups and the amount of money they're raising, the great majority of which goes to the cloud providers. Why is that? Well, the reason so much money is needed is because this is the nature of these large neural networks. They need the compute, end of story.

You can see something like this, That's all you can see a divide that's now happening between academia and the AI companies. So for a long time, for many decades, cutting edge research in AI took place in academic departments, in universities. That kept being the case up until the mid 2010s. But at some point, when the complexity and the cost of these projects started to get very large, it no longer remained possible for universities to be competitive. And now universities need, university research and AI needs to find some other way in which to contribute. Those ways exist. They're just different from the way they're used to, and different from the way the companies are contributing right now. Now, with this context, you're saying, okay, the thing about nonprofit, a nonprofit is that people who give money to a non-profit never get to see any of it back. It is a real donation. And believe it or not, it is quite a bit harder to convince people to give money to a non-profit.

And so we think what's the solution there or what is a good course of action? So we came up with an idea that to my knowledge, is unique in all corporate structures in the world. The OpenAI corporate structure is absolutely unique. OpenAI is not a for-profit company, it is a capped profit company. And I'd like to explain what that means. What that means is that equity in OpenAI can be better seen as a bond rather than equity in a normal company. And the main feature of a bond is that once it's paid out, it's gone. So in other words, OpenAI has a finite obligation to its investors, as opposed to an infinite obligation to, that normal companies have. - And does that include the founders? 00:27:06,540 Do the founders have equity in OpenAI? - So Sam Altman does not have equity, 00:27:11,940 but the other founders do. - And is it capped or is it unlimited? 00:27:15,420 - It is capped. 00:27:17,520 - And how does that cap is that capped at, 00:27:20,790 because the founders I presume didn't buy in , unless it's capped at the nominal share value? - I'm not sure I understand the question precisely, 00:27:31,140 but what I can say- - (indistinct) 00:27:33,240 - Like what, I can answer the part 00:27:35,703 which I do understand, which is like, there is certainly, like, it is, there are, it is a different, it is different from normal startup equity, but there are some similarities as well, where the earlier you join the company, the higher the cap is, because then the larger cap is needed to attract the initial investors.

As the company continues to succeed, the cap decreases. And why is that important? It's important because it means that the company, when, once all the obligation to investors and employees are paid out, OpenAI becomes a nonprofit again. And you can say this is totally crazy, what are you talking about? Like, it's not going to change anything. But it's worth considering what we expect, like, it's worth looking at what we think AI will be. I mean, we can look at what AI is today, and I think it is not at all inconceivable for open AI to achieve its, to pay out its obligation to the investors and employees, become a nonprofit, at around the time when, perhaps, the computers will become so capable, where the economic disruption will be very big, where this transition will be very beneficial. So this is the answer on the capped-profit versus nonprofit. There was a

last part of your question. I know I'm speaking for a while, but the question had many parts. The last part of your question is the Microsoft relationship. And, so here, the thing that's very fortunate, is that Microsoft is a, they're thinking about these questions the right way.

They understand the potential and the gravity of AGI. And so for example, on the, on all the investor documents that any investor in OpenAI has signed, and by the way, Microsoft is an investor into OpenAI, which is a very different relationship from the Google DeepMind. Anyone who signed any document, any investment document, there is a purple rectangle at the top of the investment document which says that the fiduciary duty of OpenAI is to the OpenAI mission, which means that you run the risk of potentially losing all your money if the mission comes in conflict. So this is something that all the investors have signed. - And lemme just make this clear for everybody, 00:30:17,580 because Google acquired DeepMind. So DeepMind was just an asset inside of Google, but beholden to Google, you're making the distinction that with OpenAI, Microsoft is an investor, and so beholden to this fiduciary duty for the mission of OpenAI, which is held by the nonprofit, which is a, is a GP or an LP in the for-profit. Okay, understood. - Yeah. Something like this. 00:30:45,303 You know, I am, you know, there are people, I can't tell you the precise details.

But so, but this is the general picture. - And you know, some have claimed though now, 00:30:59,730 especially like it, Steve Wozniak, the co-founder of Apple and Elon Musk have famously signed this very public petition saying that the point of no return is already passed, or we're approaching it, where it's gonna be impossible to rein in AI. And it's repercussions if we don't halt it now. And they've called for halting AI. I'm curious on, you are a world citizen Ilya, you were born in Russia, you were raised in Israel, you're Canadian, and I'm, and it's, open AI's response to that public petition was, I know Sam basically said that, you know, this wasn't the right way to go about doing that. But also in parallel, Sam is on a world tour with many countries that also can be antagonistic towards the west. Are there any citizen obligations, ethical obligations that you think also overweigh your technological obligations when it comes to spreading the technology around the world right now through OpenAI? Do you think that should be beholden to a regulation or some oversight? - Let me think. 00:32:17,610 Once again, the question hit a number of parts and I'd like to - It did, I apologize. 00:32:20,850 I'm trying to give you the, so you can respond however you want to on that. I know we're gonna come out of, off of time, so I just want to give you the mic, and just share everything that's on my mind, and you can decide how you wanna handle it.

- Yeah, thank you. 00:32:33,143 I mean, you know, it is true that AI is going to become truly extremely powerful, and truly extremely transformative. And I do think that we will want to move to a world with sensible government regulations. And there, you know, there are several dimensions to it. We want to be in a world where there are clear rules about, for example, training more powerful neural networks. We want there to be some kind of careful evaluation, careful prediction of these, of what we expect these neural networks, of what they can do today, and of what we expect them to be able to do, let's say, in a year from now or by the time they finish training. I think all these things will be very necessary in order to, like, rationally, I wouldn't use the word slow down the progress, I would use the term, you want to make it so that the progress is sensible. So that with each step, we've done the homework, and indeed we can make a credible story that, okay, the neural network, the system that we've trained, it has, we are doing this, and here all the steps, and it's been verified or certified. I think that is the world that we are headed to, which I think is correct. And as for the citizen obligation, I feel like, I mean, I'll answer it like this, like I think like there are two answers to it.

So obviously, you know, I live in the United States, and I really like it here, and I want this place to flourish as much as possible. I care about that. I think that of course there will be lots of, but the world is much more than just the US, and I think that these are the kind of questions which I feel are a little bit, let's say outside of my expertise, how these between country relationships work out. But I'm sure there will be lots of discussions there as well. - Okya, Ilya, can I turn a little bit towards strategy? 00:35:12,660 I'm curious for you guys internally, what metrics do you track as your north star? What are the most sacred KPIs that you use to measure OpenAI's success right now? - The most sacred KPIs? 00:35:29,580 You know, I think this is also the kind of question where maybe different people will give you different answers, but I would say that there are, if I were to really narrow it down, I would say that there are, there is a couple of really important KPI, of really important dimensions of progress. One is undeniably the technical progress. Are we doing good research? Do we understand our systems better? Are I'll be able to train them better? Can we control them better? I mean, is our research plan being executed well? Is our safety plan being executed well? How happy are we with it? I would say this would be my description of the primary KPI, which is do a good job with technology. Then there is of course stuff around the product, but, which I think is cool. But I would say that it is really the core technology which is the heart of OpenAI, the technology, its development, and its control, its steering. - And do you view right now ChatGPT as a destination? 00:36:49,500 Do you view OpenAI in the future being a destination that people go to, like Google? Or will it be powering other applications, and be the backend, or be, you know, used as part of the backend infrastructure? Is it a destination, or is it gonna be more behind the scenes in 5 to 10 years? - Yeah, well, I mean, things change so fast.

00:37:16,620 I cannot make any claims about 5 to 10 years in terms of the correct shape of the product. I imagine a little bit of both perhaps, but this kind of question, I mean, I think it remains to be seen. I think there are, I think this stuff is still so new. - Okay, I'm gonna ask one more question, 00:37:36,180 then I'm gonna jump to the student questions. If you were a student at Stanford today interested in AI, if you were, you know, somebody who wants to be Ilya, what would you focus your time? And another second question on this. If you're also interested in entrepreneurship, where would, what would you, what advice would you give for a Stanford undergrad engineer that's interested in AI and entrepreneurship? - So I think on the first

one, 00:38:07,500 it's always hard to give generic advice like this. - Yeah. 00:38:12,843 - But, I can still provide some generic advice nonetheless. 00:38:17,523 And I think it's something like, it is generally a good idea to lean into one's unique predispositions. You know, every, why, if you think, if you look, if you think about the set of, let's say, inclinations or skills or talents that the person may have, the combination is pretty rare.

So leaning into that is a very good idea, no matter which direction you choose to go, look to go in. And then on the AI research, like I would say that there, you know, I could say something, but even, but there especially, you want to lean into your own ideas, and really ask yourself what, can you, is there something that's totally obvious to you that makes you go, why is everyone else not getting it? If you feel like this, that's a good sign. It means that you might be able, that you want to lean into that, and explore it and see if your instinct is true or not. It may not be true, but you know, my advisor Geoff Hinton says this thing which I really like. He says, you should trust your intuition. Cause if your intuition is good, you go really far, and if it's not good, then there's nothing you can do. And as far as entrepreneurship is concerned, I feel like this is a place where the unique perspective is even more valuable. Or maybe equally, it's because it's, maybe I'll explain why I think it's more valuable than in research, well, in research it's very valuable too, but in entrepreneurship, like you need to, like, almost pull from your unique life experience. Where you say, okay, I see this thing, I see this technology, I see some, like, take a very, very broad view, and see if you can hone in on something, and then actually just go for it. So that would be the conclusion of my generic advice.

- Okay, which is great. 00:40:26,040 That's also great. I'm gonna move on to the student question. So one of the most upvoted question is how do you see the field of deep learning evolving in the next 5 to 10 years? - Yeah, let's see. 00:40:41,460 You know, I expect deep learning to continue to make progress. I expect that, you know, there was a period of time where a lot of progress came from scaling, and you, we saw that most- in the most pronounced way in going from GPT-1 to GPT-3. But things will change a little bit. The reason that progress in scaling was so rapid is because people had all these data centers, which they weren't using for a single training run. So by simply reallocating existing resources, you could make a lot of progress. And it doesn't take that long necessarily to reallocate existing resources.

You just need to, you know, someone just needs to decide to do so. It is different now because the training runs are very big, and the scaling is not going to be progressing as fast as it used to be, because building data center takes time. But at the same time, I expect deep learning to continue to make progress in (indistinct) from other places. The deep learning stack is quite deep, and I expect that there will be improvements in many layers of the stack. And together they will still lead to progress being very robust. And so, if I had to guess, I'd imagine that there would be maybe, I'm certain we will discover new properties which are currently unknown of deep learning, and those properties will be utilized. And I fully expect that the systems of 5 to 10 years from now will be much, much better than the ones they are, we have right now. But exactly how it's going to look like, I think it's a bit harder to answer. It's a bit like, it's because the improvements that there is, there will be, maybe small number of big improvements, and also a large number of small improvements, all integrated into a large complex engineering artifact. - And can I ask your, you know, 00:42:46,560 your co-founder Sam Altman has said that we've reached the limits of what we can achieve by scaling to larger language models.

Is, do you agree, and if so, you know, what, then what is the next innovation frontier that you're focusing on, if that's the case? - Yeah, so 00:43:06,150 I think maybe, I don't remember- I don't know exactly what he said, but maybe he meant something like that the age of easy scaling has ended or something like this. Like of course the larger neural nets will be better, but it's will be a lot of effort and cost to do them. But I think there will be lots of different frontiers. And actually, to the question of how can one contribute in deep learning, identifying such a frontier, perhaps one that's been missed by others, is very fruitful. - And is it, can I go even just deeper on that, 00:43:38,850 because I think there is this debate about vertical focus versus generalist training. You know, is it better, do you think the- there's better performance that can be achieved in particular domains such as law or medicine by training with special data sets? Or is it likely that generalist training with all available data will be more beneficial? - So, like at some point, 00:44:02,580 we should absolutely expect specialist training to make huge impact. But the reason we do the generalist training is just so that we can even reach the point where- just so that we can reach the point where the neural network can even understand the questions that we are asking. And only when it has a very robust understanding, only then we can go into specialist training and really benefit from it. So yeah, I mean I think all these, I think these are all fruitful directions. - But you don't think, 00:44:36,300 when do you think we'll be at that point, when specialist training is the thing to focus on? - I mean, 00:44:46,170 you know, like if you look at people who do open source work, people who work with open open source models, they do a fair bit of this kind of specialist training, because they have a fairly underpowered model and they try to get any ounce of performance they can out of it.

So I would say that this is an example. I'd say that this is an example of it happening. Like it's already happening to some degree. It's not a binary, it's a, you might wanna think of it as of a, like a continuum, a spectrum. - But do you think that the competitive, 00:45:19,890 do you think that the winning advantage is gonna be having these proprietary data sets, or is it gonna be having a much higher performance, large language model when it comes to these applications of AI into verticals? - So I think it's maybe productive to think 00:45:37,290 about an AI like this as a combination of multiple factors, where each factor makes a contribution. And is it better to have special data, which helps you make your AI better in a particular set of tasks, of course. Is it better to have a more capable base model, of course, from the perspective of the task. So maybe this is the answer, it's not an either or. - Okay, I'm gonna move down to the other questions. 00:46:04,530 There's a question on

what was the cost of training and developing GPT-3/4? - Yeah, so, 00:46:14,250 you know, for obvious reasons, I can't comment on that.

- Okay, but there, I think there is a, 00:46:21,180 you know, I think even from our research community, there's a strong desire to be able to get access to different aspects of OpenAI's technology. And are there any plans for releasing it to researchers, or to other startups to encourage more competition and innovation? Some of the requests that I've heard are unfettered interactions without safeguards to understand the model's performance model specifications, including details on how it was trained, and access to the model itself, i.e. the trained parameters. Do you wanna comment on any of that? - Yeah, I mean, I think 00:47:05,100 like it's related to our earlier question about open versus closed, I think that there are some intermediate approaches which can be very fruitful. For example, model access and various combinations of that can be very, very productive, because these neural networks already have such a large and complicated surface area of behavior. And studying that alone can be extremely interesting. Like, we have an academic access program, we provide various forms of access to the models. And in fact, plenty of academic research labs do study them in this way. So I think this kind of approach is viable and it's something that we could, that we are doing. - I know we're coming up on time.

00:48:01,860 I wanna just end with just one final question, which is can you just share any unintuitive but compelling use cases for how you love to use ChatGPT that others may not know about? - So, I mean, I don't, I wouldn't say that it's unknown, 00:48:24,003 but I, I really enjoy its poem writing ability. It can write poems, it can rap, it can, it can be pretty amusing. - And do you guys use it? 00:48:38,053 Is it an integrated part of the, of teamwork at Open? I assume it is, but I'm curious, do you have any insights on how it changes dynamics with teams when you have AI deeply integrated into, you know, a human team and how they're working and any insights into what we may not know, but that will come? - I would say today, the best way 00:48:59,110 to describe the impact is that everyone is a little bit more productive. People are a little bit more on top of things. I wouldn't say that right now there is a dramatic impact on dynamics, which I can say, oh yeah, the dynamics have shifted in this pronounced way. - Okay, I'm curious if it depersonalizes conversations 00:49:18,300 because it's the AI bot, or maybe it may, but maybe we're not at that point yet where it's becoming the same- - I 00:49:26,610 that I definitely, I don't think that's the case, and I predict that'll not be the case, but we'll see. - Okay, well thank you Ilya for a fascinating discussion. 00:49:36,840 Time is always too short. You're always invited back to The Farm. We'd love to have you, either virtually or in person.

So thank you, thank you, thank you. To our audience, thank you for tuning in for this session of the Entrepreneurial Thought Leader Series. Next week we'll be gonna be joined by the executive chairman and co-founder of Okta, Frederic Kerrest. And you can find that event, and other future events in this ETL series on our Stanford eCorner YouTube channel. And you'll find even more of the videos, podcasts, and articles about entrepreneurship and innovation at Stanford eCorner, that's ecorner.stanford.edu. And as always, thank you for tuning in to ETL. (electronic music)..