URL: https://ecorner.stanford.edu/videos/the-possibilities-of-ai-entire-talk/

Sam Altman is the co-founder and CEO of OpenAI, the AI research and deployment company behind ChatGPT and DALL-E. Altman was president of the early-stage startup accelerator Y Combinator from 2014 to 2019. In 2015, he co-founded OpenAI as a nonprofit research lab with the mission to build general-purpose artificial intelligence that benefits all humanity. In this conversation with Stanford adjunct lecturer Ravi Belani, Altman gives advice for aspiring AI entrepreneurs and shares his insights about the opportunities and risks of AI tools and artificial general intelligence.

Transcript

(upbeat music) - Welcome to the Entrepreneurial Thought Leader Seminar 00:00:20,783 at Stanford University. (audience applauding) This is the Stanford Seminar for Aspiring Entrepreneurs. ETL is brought to you by STVP, the Stanford Entrepreneurship Engineering Center and BASES, the Business Association of Stanford Entrepreneurial Students. I'm Ravi Belani, a lecturer in the Management Science and Engineering Department, and the director of Alchemist and Accelerator for Enterprise Startups. And today I have the pleasure of welcoming Sam Altman to ETL. (audience applauding) Sam is the co-founder and CEO of OpenAI. OpenAI is not a word I would use to describe the seats in this class. And so I think by virtue of that, that everybody already probably knows OpenAI, but for those who don't, OpenAI is the research and deployment company behind ChatGPT, DALL-E and Sora. Sam's life is a pattern of breaking boundaries and transcending what's possible both for himself and for the world. He grew up in the Midwest in St.

Louis, came to Stanford, took ETL as an undergrad and we held onto Stanford for Sam for two years he studied computer science and then after his sophomore year he joined the inaugural class of Y Combinator with a social mobile app company called Loopt that then went on to go raise money from Sequoia and others. He then dropped out of Stanford, spent seven years on Loopt, which got acquired and then he rejoined Y Combinator in an operational role. He became the president of Y Combinator from 2014 to 2019. And then in 2015 he co-founded OpenAI as a nonprofit research lab with the mission to build general purpose artificial intelligence that benefits all humanity. OpenAI has set the record for the fastest growing app in history with the launch of ChatGPT, which grew to 100 million active users just two months after launch. Sam was named one of Time's 100 most influential people in the world. It was also named Time's CEO of the year in 2023. And he was also most recently added to Forbes list of the World's Billionaires. Sam lives with his husband in San Francisco and splits his time between San Francisco and Napa and he's also a vegetarian. And so with that, (audience laughing) please join me in welcoming Sam Altman to the stage.

(audience applauding) And in full disclosure, that was a longer introduction than Sam probably would've liked. Brevity is the soul of wit and so we'll try to make the questions more concise. - Great. 00:02:47,310 this is also Sam's birth week. It was his birthday on Monday and I mentioned that just because I think this is an auspicious moment both in terms of time, you're 39 now, and also place, you're at Stanford in ETL that I would be remiss if this wasn't sort of a moment of just some reflection. And I'm curious if you reflect back on when you were half a life younger, when you were 19 in ETL, if there were

three words to describe what your felt sense was like as a Stanford undergrad, what would those three words be? - It's always hard questions. 00:03:20,310 I was like, ex, you want three words only? Okay. (audience laughing) - You can, 00:03:23,880 you can go more Sam. You're the king of brevity. - Excited, optimistic, and curious.

00:03:31,770 - Okay. And what would be your three words now? 00:03:32,910 - I guess the same. 00:03:33,750 - Which is terrific. 00:03:35,070 So there's been a constant thread even though the world has changed. And you know, a lot has changed in the last 19 years, but that's gonna pale in comparison to what's gonna happen in the next 19. - Yeah. 00:03:45,150 - And so I need to ask you for your advice 00:03:46,740 if you were a Stanford undergrad today. So if you had a freaky Friday moment tomorrow you wake up and suddenly you're 19 in inside of as a Stanford undergrad, knowing everything you know, what would you do? Would you drop out? - I'd be very happy. 00:03:58,920 I would feel like I was like coming of age at the luckiest time, like in several centuries probably. I think the degree to which the world is is gonna change and the opportunity to impact that, starting a company, doing AI research, any number of things is like quite remarkable.

I think this is probably the best time to start. Yeah, I think I would say this, I think this is probably the best time to start a company since the internet at least. And maybe kind of like in the history of technology. I think with what you can do with AI is, like gonna just get more remarkable every year. And the greatest companies get created at times like this. The most impactful new products get built at times like this. So I would feel incredibly lucky and I would be determined to make most of it and I would go figure out like where I wanted to contribute and do it. - And do you have a bias on where would you contribute? 00:04:56,010 Would you wanna stay as a student? Would and if so would you major in a certain major giving the pace of change? - Probably I would not stay as a student 00:05:03,960 but only 'cause like I didn't, and I think it's like reasonable to assume people kind of are gonna make the same decisions they would make again. I think staying as a student is a perfectly good thing to do. It would probably not be what I would have picked.

- And this is you, this is you. 00:05:18,660 So you have the freaky Friday moment, it's you, you're reborn and as a 19-year-old and what would you? Yeah? - What I think I would, 00:05:27,270 again like I think this is not a surprise 'cause people kind of are gonna do what they're gonna do. I think I would go work on AI research. (audience laughing) - And where might you do that Sam? 00:05:37,590 - I think, I mean obviously I have a bias towards OpenAI 00:05:39,960 but I think anywhere I could like do meaningful AI research, I would be like very thrilled about. - So you'd be agnostic 00:05:44,280 if that's academia or private industry. - I say this with sadness, 00:05:52,410 I think I would pick industry, realistically. I think it's, I think you kind of need to be at a place with so much compute. - Okay. 00:06:03,120 And if you did join on the research side, would you join, so we had Qasar here last week who was a big advocate of not being a founder but actually joining an existing company, sort of learn the chops. For the students that are wrestling with should I start a company now at 19 or 20 or should I go join another entrepreneur, either research lab or venture? What advice would you give them? - Well since he gave the case to join a company, 00:06:28,470 I'll give the other one which is I think you learn a lot just starting a company and if that's something you want to do at some point there's this thing Paul Graham says, but I think it's like very deeply true.

"There's no pre-startup like there is pre-med." You kind of just learn how to run a startup by running a startup. And if that's what you're pretty sure you wanna do, you may as well jump in and do it. - And so let's say, 00:06:46,745 so if somebody wants to start a company and they want to be in AI, what do you think are the biggest near term challenges that you're seeing in AI that are the ripest for a start startup? And just to scope that, what I mean by that are what are the holes that you think are the top priority needs for OpenAI that OpenAI will not solve in the next three years? Yeah. - So, 00:07:11,790 I think this is like a very reasonable question to ask in some sense, but I think it's, I'm not gonna answer it because I think you should never take this kind of advice about what startup to start ever from anyone. I think by the time there's something that is like the kind of thing that's obvious enough that me or somebody else will sit up here and say it, it's probably like not that great of a startup idea. And I totally understand the impulse and I remember when I was just like asking people like what startup should I start? But I think like one of the most important things I believe about having an impactful career is you have to chart your own course. If the thing that you are thinking about is something that someone else is gonna do anyway or more likely something that a lot of people are gonna do anyway, you should be like somewhat skeptical of that. And I think a really good muscle to build is coming up with the ideas that are not the obvious ones to say. So I don't know what the really important idea is that I'm not thinking of right now, but I'm very sure someone in this room does it, knows what that answer is. And I think learning to trust yourself and come up with your own ideas and do the very like non-consensus things like when we started OpenAI, that was an extremely non-consensus thing to do and now it's like the very obvious thing to do.

Now I only have the obvious ideas 'cause I'm just like stuck in this one frame, but I'm sure you all have the other ones. - But so can I ask it another way? 00:08:43,110 And I don't know if this is fair or not, but what questions then are you wrestling with that no one else is talking about? - How to build really big computers? 00:08:53,250 I mean I think other people are talking about that, but we're probably like looking at it through a lens that no one else is quite imagining yet. I mean we're definitely wrestling with how we, when we make not just like grade school or middle school or level intelligence but like PhD level intelligence and beyond the best way to put that into a product, the best way to have a positive impact with that on society and people's lives. We don't know the answer to that yet. So I think that's like a pretty important thing to figure out. - Okay. 00:09:25,980 And can we continue on that thread then of how to build really big computers if that's really what's on your mind. Can you share, I know there's been a lot of speculation and that probably a lot of hearsay too about the semiconductor foundry endeavor that you are reportedly embarking on. Can you share what would make, what's the vision? - Yeah- 00:10:04,230 And that is energy, data centers, chips, chip design, new kinds of networks. It's how we look at that entire

ecosystem and how we make a lot more of that.

And I don't think it'll work to just look at one piece or another, but we, we gotta do the whole thing. - Okay, so there's multiple big problems. 00:10:20,880 - Yeah. 00:10:23,190 I think like just this is the arc of human technological history as we build bigger and more complex systems. - And does it grow? 00:10:31,830 So, you know, in terms of just like the compute cost, correct me if I'm wrong, but ChatGPT 3 was, I've heard it was $100 million to do the model and it was 175 billion parameters. GPT 4 was cost $400 million with 10X the parameters, it was almost 4X the cost. But 10X the parameters. Correct me, adjust me- - You know it, 00:10:57,513 This is Stanford Sam. Okay. But even if you don't want to correct the actual numbers, if that's directionally correct, does the cost do you think keep growing with each subsequent? - Yes.

00:11:09,300 - Yes. 00:11:11,400 And does it keep growing multiplicatively? - Probably. I mean. (audience laughing) - And so then the question then becomes how do we, 00:11:20,403 how do you capitalize that? - Well look, 00:11:24,640 I kind of think that giving people really capable tools and letting them figure out how they're gonna use this to build the future is a super good thing to do and is super valuable and I'm super willing to bet on the ingenuity of you all and everybody else in the world to figure out what to do about this. So there is probably some more business-minded person than me at OpenAI somewhere that is worried about how much we're spending, but I kind of don't. (audience laughing) - Okay so that doesn't cross it. 00:12:00,630 So, you know, OpenAI is phenomenal, ChatGPT is phenomenal, everything else, all the other models are phenomenal. You've burned $520 million of cash last year that doesn't concern you in terms of thinking about the economic model of how do you actually, where's gonna be the monetization source? - Well, first of all, that's nice of you to say, 00:12:17,670 but ChatGPT is not phenomenal. Like ChatGPT is like mildly embarrassing at best. GPT 4 is the dumbest model any of you will ever, ever have to use again by a lot.

But, you know, it's like important to ship early and often and we believe in iterative deployment. Like if we go build AGI in a basement and then, you know, the world is like kind of blissfully walking blindfolded along. I don't think that's like, I don't think that makes us like very good neighbors. So I think it's important given what we believe is going to happen to or express our view about what we believe is gonna happen. But more than that, the way to do it is to put the product in people's hands and let society co-evolve with the technology. Let society tell us what it collectively and people individually want from the technology. How to productize this in a way that's gonna be useful. Where the model works really well, where it doesn't work really well. Give our leaders and institutions time to react, give people time to figure out how to integrate this into their lives to learn how to use the tool for some of you all like cheat on your homework with it. But some of you all probably do like very amazing, wonderful things with it too.

And as each generation goes on, I think that will expand and that means that we ship imperfect products but we have a very tight feedback loop and we learn and we get better and it does kind of suck to ship a product that you're embarrassed about but it's much better than the alternative. And in this case, in particular, where I think we really owe it to society to deploy iteratively, one thing we've learned is that AI and surprise don't go well together. People don't wanna be surprised, people want a gradual rollout and the ability to influence these systems. That's how we're gonna do it. And there may be, there could totally be things in the future that would change where we think iterative deployment isn't such a good strategy, but it does feel like the current best approach that we have and I think we've gained a lot from doing this and, you know, hopefully the larger world has gained something too. Whether we burn 500 million a year or 5 billion or 50 billion a year, I don't care. I genuinely don't. As long as we can, I think, stay on a trajectory where eventually we create way more value for society than that and as long as we can figure out a way to pay the bills like we're making AGI, it's gonna be expensive. It's totally worth it. - And so do you have a, I hear you.

00:14:58,380 Do you have a vision in 2030 of what if I say you crushed it Sam, it's 2030, you crushed it. What does the world look like to you? - You know, maybe in some very important ways 00:15:09,993 not that different. Like we will be back here, there will be like a new set of students. We'll be talking about how startups are really important. Technology is really cool. We'll have this new great tool in the world, it'll feel, it would feel amazing if we got to teleport forward six years today and have this thing that was like smarter than humans in many subjects and could do these complicated tasks for us. And, you know, like we could have these like complicated program written or this research done or this business started and yet like the sun keeps rising, like people keep having their human dramas, life goes on. So sort of like super different in some sense that we now have like abundant intelligence at our fingertips and then in some other sense, like not different at all. - Okay. 00:16:05,190 And you mentioned artificial general intelligence, AGI, artificial general intelligence.

And in a previous interview you defined that as software that could mimic the median competence of a or the competence of a median human for tasks. - Yeah. 00:16:18,000 - Can you gimme, is there a time, 00:16:19,890 if you had to do a best guess of when you think or a range- - You know, I- 00:16:22,200 - When you feel like that's gonna happen. 00:16:25,500 - I think we need a more precise definition of AGI 00:16:26,830 for the timing question because at this point, even with like the definition you just gave, which is a reasonable one, there's, so that's- - I'm parroting back what you said in an interview. 00:16:37,530 - Well that's good 'cause I'm gonna criticize myself. 00:16:38,365 - Okay. (audience laughing) - It's too loose of a definition. 00:16:44,940 There's too much room for misinterpretation in there to I think be really useful or get at what people really want. Like I kind of think what people wanna know when they say like what's the timeline to AGI is like when is the world gonna be super different? When is the rate of change gonna get super high? When is the way the economy works gonna be really different? Like when does my life change? And that for a bunch of reasons may be very different than we think. Like

I can totally imagine a world where we build PhD level intelligence in any area and, you know, we can make researchers way more productive.

Maybe we can even do some autonomous research and in some sense like that sounds like it should change the world a lot and I can imagine that we do that and then we can detect no change in global GDP growth for like years afterwards, something like that. Which is very strange to think about. And it was not my original intuition of how this was all gonna go. So I don't know how to give a precise timeline of when we get to the milestone people care about, but when we get to systems that are way more capable than we have right now one year and every year after. And that I think is the important point. So I've given up on trying to give the AGI timeline, but I think every year for the next many, we have dramatically more capable systems every year. - I wanna ask about the dangers of AGI. 00:18:12,030 And gang, I know there's tons of questions for Sam. In a few moments I'll be turning it up. So start start thinking about your questions.

A big focus on Stanford right now is ethics and can we talk about, you know, how you perceive the dangers of AGI and specifically do you think the biggest danger from AGI is gonna come from a cataclysmic event which, you know, makes all the papers or is it gonna be more subtle and pernicious sort of like, you know, like how everybody has ADD right now from, you know, using TikTok. - Right. 00:18:39,030 - Are you more concerned about the subtle dangers 00:18:42,390 or the cataclysmic dangers or neither? - I'm more concerned about the subtle dangers 00:18:46,680 because I think we're more likely to overlook those. The cataclysmic dangers, a lot of people talk about and a lot of people think about and I don't wanna minimize those, I think they're really serious and a real thing. But I think we, at least, know to look out for that and spend a lot of effort. The example you gave of everybody getting ADD from TikTok or whatever, I don't think we knew to look out for and that that's a really hard, the unknown unknowns are really hard and so I'd worry more about those, although I worry about both. - And are they, unknown unknowns, 00:19:19,770 are there any that you can name that you're particularly worried about? - Well then I would kind of, they'd be the known unknown. 00:19:25,493 (audience laughing) - But you can. 00:19:30,720 - I am worried just about so, 00:19:32,430 so even though I think in the short term things change less than we think as with other major technologies, in the long term I think they change more than we think. And I am worried about what rate society can adapt to something so new and how long it'll take us to figure out the new social contract versus how long we get to do it.

I'm worried about that. - Okay. 00:19:55,480 I'm gonna open up soon. I just wanna ask you a question about one of the key things that we're now trying to inculcate into the curriculum as things change so rapidly is resilience. - Hmm. That's really good. 00:20:06,570 - And, you know, for, (audience laughing) and the cornerstone of resilience is self-awareness. And I'm wondering if you feel that you're pretty self-aware of your driving motivations as you are embarking on this journey? - So first of all I think, 00:20:25,170 I believe resilience can be taught, I believe it has long been one of the most important life skills and in the future I think in the, over the next couple of decades, I think resilience and adaptability will be more important than they've been in a very long time. So I think that's really great. On the self-awareness question, I think I'm self aware but I think like everybody thinks they're self-aware and whether I am or not is sort of like hard to say from the inside.

- And can I ask you sort of the questions 00:20:58,320 that we ask in our intro classes on self-awareness? - Sure. 00:21:22,623 you can seek connections across them. I think you can then kind of come up with the ideas that are different than everybody else has or that's sort of the experts in one area have. - And what are your most dangerous weaknesses? 00:21:40,493 - Most dangerous, that's an interesting framework for it. 00:21:45,030 I think I have like a general bias to be too pro-technology just 'cause I'm curious and I wanna see where it goes. And I believe that technology is on the whole, a net good thing. But I think that is a worldview that has overall served me and others well and thus gotten like a lot of positive reinforcement and is not always true and when it's not been true has been like pretty bad for a lot of people. - And then Harvard psychologist David McClelland 00:22:11,280 has this framework that all leaders are driven by one of three primal needs: a need for affiliation, which is a need to be liked, a need for achievement and the need for power. If you had to rank list those, what would be yours? - I think at various times in my career, all of those, 00:22:28,913 I think they're these like levels that people go through. At this point I feel driven by like wanting to do something useful and interesting.

- Okay. 00:22:37,546 - But I think- 00:22:41,790 - Okay, and then where were you when you most last 00:22:43,443 felt most like yourself? - I always feel like. (audience laughing) - Okay, and then one last question I'll 00:22:53,250 and what are you most excited about with ChatGPT 5 that's coming out that people don't, what are you most excited about with the release of ChatGPT 5 that we're all gonna see? - I don't know yet. 00:23:07,110 I mean I, this sounds like a cop out answer but I think the most important thing about GPT 5 or whatever we call that is just that it's gonna be smarter and this sounds like a dodge but I think that's like among the most remarkable facts in human history that we can just do something and we can say right now with a high degree of scientific certainty, GPT 5 is gonna be smarter than, a lot smarter than GPT 4, GPT 6 can be a lot smarter than GPT 5 and we are not near the top of this curve and we kind of know what to do and this is not like it's gonna get better in one area. This is not like we're gonna, you know, it's not that it's always gonna get better at this eval or this subject or this modality. It's just gonna be smarter in the general sense. And I think the gravity of that statement is still like underrated. - Okay, that's great. 00:23:52,890 Guys, Sam is really here for you. - Yeah.

00:24:02,130 I sort of wanted to talk to you about responsible deployment of AGI. So as you guys continually inch closer to that, how do you plan to deploy that responsibly at OpenAI, you know, to prevent, you know, stifling human innovation and continue to spur that? - So I'm actually not worried at all about stifling 00:24:22,073 of human innovation. I really deeply believe that people will just surprise us on the upside with better tools. I think all of history suggests that if you give people

more leverage, they do more amazing things and that's kind of like we all get to benefit from that. That's just kind of great. I am though increasingly worried about how we're gonna do this all responsibly. I think as the models get more capable, we have a higher and higher bar. We do a lot of things like red teaming and external audits and I think those are all really good but I think as the models get more capable, we'll have to deploy even more iteratively, have an even tighter feedback loop on looking at how they're used and where they work and where they don't work. And this world that we used to do where we can release a major model update every couple of years, we probably have to find ways to like increase the granularity on that and deploy more iterative than we have in the past. And it's not super obvious to us yet how to do that.

But I think that'll be key to responsible deployment and also the way we kind of have all of the stakeholders negotiate what the rules of AI need to be. That's gonna get more complex over time too. - Thank you. 00:25:32,940 Next question over here. Woman You mentioned before that there's a growing need 00:25:38,280 for larger and larger computers and faster computers. However, many parts of the world don't have the infrastructure to build those data centers or those large computers. How do you see global innovation being impacted by that? - So two parts to that. 00:25:53,220 One, no matter where the computers are built, I think global and equitable access to use the computers for training as well as inference is super important. One of the things that's like very core to our mission is that we make ChatGPT available for free to as many people as want to use it with the exception of certain countries where we either can't or don't for a good reason wanna operate. How we think about making training compute more available to the world is gonna become increasingly important.

I do think we get to a world where we sort of think about it as a human right to get access to a certain amount of compute and we gotta figure out how to like distribute that to people all around the world. There's a second thing though, which is I think countries are going to increasingly realize the importance of having their own AI infrastructure and we want to figure out a way and we're now spending a lot of time traveling around the world to build them in the many countries that'll wanna build these and I hope we can play some small role there in helping that happen. - Terrific, thank you. 00:26:56,040 Man My question was what role do you envision for AI 00:26:58,380 in the future of like space exploration or like colonization? - I think space is like not that hospitable 00:27:06,210 for biological life obviously. And so if we can send the robots that seems easier. (audience laughing) Woman Hey Sam, so my question is 00:27:21,480 for a lot of the founders in the room and I'm gonna give you the question and then I'm gonna explain why I think it's complicated. So my question is about how you know an idea is non-consensus and the reason I think it's complicated is 'cause it's easy to overthink. I think today even yourself says AI is the place to start a company. I think that's pretty consensus, maybe rightfully so. It's an inflection point.

I think it's hard to know if a idea is non-consensus depending on the group that you're talking about. The general public has a different view of tech from the tech community and even tech elites have a different point of view from the tech community. So I was wondering how you verify that your idea is non-consensus enough to pursue. - I mean first of all, 00:28:12,690 what you really want is to be right. Being contrarian and wrong is still is wrong. And if you predicted like 17 outta the last two recessions, you probably were contrarian for the two you got right, probably not even necessarily. But you were wrong 15 other times. And so I think it's easy to get too excited about being contrarian and again like the most important thing to be right and the group is usually right, but where the most value is is when you are contrarian and right. And that doesn't always happen in like sort of a 0 or 1 kind of way. Like everybody in the room can agree that AI is the right place to start a company and if one person in the room figures out the right company to start and then successfully executes on that and everybody else thinks ah, that wasn't the best thing you could do.

That's what matters. So it's okay to kind of like go with conventional wisdom when it's right and then find the area where you have some unique insight. In terms of how to do that I do think surrounding yourself with the right peer group is really important and finding original thinkers is important. But there is part of this where you kind of have to do it solo or at least part of it solo or with a few other people who are like, you know, gonna be your co-founders or whatever. And I think by the time you're too far in the like, how can I find the right peer group? You're somehow in the wrong framework already. So like learning to trust yourself and your own intuition and your own thought process, which gets much easier over time. No one, no matter what they say, I think is like truly great at this when they're just starting out. Yeah, because like you kind of just haven't built the muscle and like all of your social pressure and all of like the evolutionary pressure that produced you was against that. So it's something that like you get better at over time and don't hold yourself to too high of a standard too early on it. Man Hi Sam.

00:30:23,430 I'm curious to know what your predictions are for how energy demand will change in the coming decades and how we achieve a future where renewable energy sources are 1 cent per kilowatt hour. - I mean it will go up for sure. 00:30:36,983 Well, not for sure. You can come up with all these weird ways in which like we all, depressing future's where it doesn't go up. I would like it to go up a lot. I hope that we hold ourselves to a high enough standard where it does go up. I forget exactly what the kind of world's electrical generating capacity is right now, but let's say it's like 3,000, 4,000 gigawatts, something like that. Even if we add another 100 gigawatts for AI, it doesn't materially change it that much, but it changes it some. And if we start to add 1,000 gigawatts for AI, someday it does, that's a material change. But there are a lot of other things that we wanna do.

And energy does seem to correlate quite a lot with quality of life we can deliver for people. My guess is that fusion eventually dominates electrical generation on earth. I think it should be the cheapest, most abundant, most reliable densest source. I could be wrong on that and it could be solar plus storage and, you know, my guess most likely is it's gonna be 80 20

one way or the other and there'll be some cases where one of those is better than the other. But those kind of seem like the two bets for like really global scale 1 cent per kilowatt hour energy. Woman Hi Sam, I have a question. 00:31:56,820 It's about OpenAI what happened last year. So what's the lesson you learned? 'Cause you talked about resilience, so what's the lesson you learned from left that company and now coming back and what made you coming back? Because Microsoft also give you offer like- - Yeah. 00:32:19,080 that totally could have run the company without me and did for a couple of days. (audience laughing) And you never, and also that the team was super resilient.

Like we knew that some crazy things and probably more crazy things will happen to us between here and AGI as different parts of the world have stronger and stronger emotional reactions and the stakes keep ratcheting it up and, you know, I thought that the team would do well under a lot of pressure, but you never really know until you get to run the experiment. And we got to run the experiment and I learned that the team was super resilient and like ready to kind of run the company. In terms of why I came back, you know, I originally when the, so it was like the next morning the board called me and were like, what do you think about coming back? And I was like, no, I'm mad. (audience laughing) And then I thought about it and I realized just like how much I loved OpenAI, how much I loved the people, the culture we had built, the mission and I kind of like wanted to finish it altogether. - You had emotionally. 00:33:27,620 I just wanted, this is obviously a really sensitive, oh it's not? But was I imagine that was most, (audience laughing) okay, Well, then could we talk about this structure about it because this Russian doll structure of the OpenAI where you have the nonprofit owning the for-profit, you know, when we're trying to teach principal driven entrepreneurship - Yeah we got here. 00:33:46,200 We got to the structure gradually. It's not what I would go back and pick if we could do it all over again. But we didn't think we were gonna have a product when we started. We were just gonna be like an AI research lab wasn't even clear.

We had no idea about a language model or an API or ChatGPT. So if you're gonna start a company, you gotta have like some theory that you're gonna sell a product someday. And we didn't think we were going to. We didn't realize we were gonna need so much money for compute. We didn't realize we were gonna like have this nice business. - So what was your intention when you started it? 00:34:15,248 - We just wanted to like push AI research forward. 00:34:16,437 - Okay. 00:34:27,663 - I cannot overstate how foreign of a concept like- 00:34:29,820 - I mean for you personally, 00:34:32,040 not for OpenAI, but you weren't starving. - Well, I had already made a lot of money 00:34:34,920 so it was not like a big, (audience laughing) I mean I like, I don't wanna like claim some like moral purity here. It was just like that was the phase of my life.

- That's not a driver? 00:34:44,216 - Not a big driver. 00:34:45,049 - Okay. 00:34:45,882 Because there's this, so, and the reason why I'm asking is just, you know, when we're teaching about principle driven entrepreneurship here, you can understand principles inferred from organizational structures. When the United States was set up, the architecture of governance is the constitution. It's got three branches of government, all these checks and balances and you can infer certain principles that, you know, there's a skepticism on centralizing power that, you know, things will move slowly. It's hard to get things to change, but it'll be very, very stable. If you, you know, not to parrot Billie Eilish, but if you look at the OpenAI structure and you think, what was that made for? It's, you're like, you're near 100 billion dollar valuation and you've got a very, very limited board that's a nonprofit board which is supposed to look after it's fiduciary duties to this- - Yeah, again 00:35:30,120 it's not what we would've done if we knew now then what we know now. But you don't get to like play life in reverse and you have to just like adapt. There's a mission we really cared about. We thought AI was gonna be really important.

We thought we had an algorithm that learned, we knew it got better with scale. We didn't know how predictably it got better with scale and we wanted to push on this. We thought this was like gonna be a very important thing in human history and we didn't get everything right, but we were right on the big stuff and our mission hasn't changed and we've adapted the structure as we go and we'll adapt it more in the future. But, you know, like you don't, like life is not a problem set. You don't get to like solve everything really nicely all at once. It doesn't work quite like it works in the classroom as you're doing it. And my advice is just like trust yourself to adapt as you go. It'll be a little bit messy but you can do it. - And I just ask this because of the significance 00:36:24,330 of OpenAI, you have a board which is all supposed to be independent financially so that they're making these decisions as a nonprofit, thinking about the stakeholder, they're stakeholder that they're fiduciary of, isn't the shareholders, it's humanity. Everybody's independent.

There's no financial incentive that anybody has that's on the board, including yourself with OpenAI. - Well Greg was, 00:36:45,330 okay first of all, I think making money is a good thing. I think capitalism is a good thing. My co-founders on the board have had financial interests and I've never once seen them not take the gravity of the mission seriously. But, you know, we've put a structure in place that we think is a way to get incentives aligned and I do believe incentives are superpowers, but I'm sure will evolve it more over time. And I think that's good, not bad. - And with OpenAI then you fund you're not, 00:37:12,570 you don't get any carry in that and you're not following on investments onto those countries. - Okay. 00:37:16,463 - Okay. Thank you.

00:37:17,296 - We can keep talking about this. 00:37:18,129 - No, I know, I know, I know. 00:37:18,962 You wanna go back to students. I do too. So we'll go. We'll keep going to the students. Man How do you expect that AGI will change geopolitics 00:37:25,710 in the balance of power in the world. - Phew. 00:37:32,720 Like maybe more than any other technology? I don't, I think about that so much and I have such a hard time saying what it's actually gonna do. I, or maybe more accurately, I have such a hard time saying what it won't do.

And we were talking earlier about how it's like not gonna change, maybe it won't change day-to-day life that much, but the balance of power in the world, it feels like it does change a lot. But I don't have a deep answer of exactly how. Woman Thanks so much. 00:38:02,100 I was wondering, sorry, I was wondering in the deployment of like general intelligence and also responsible AI, how much do you think is it necessary that AI systems are somehow capable of recognizing their own insecurities or like uncertainties and actually communicating them to the outside world? - I always get nervous anthropomorphizing AI too much 00:38:24,780 'cause I think it like can lead to a bunch of weird oversights. But if we say like how much can AI recognize its own flaws? I think that's very important to build and right now and the ability to like recognize an error in reasoning and have some sort of like introspection ability like that that that seems to me like really important to pursue. Man Hey Sam. 00:38:55,140 Thank you for giving us some of your time today and coming to speak. From the outside looking in, we all hear about the culture and togetherness of OpenAI in addition to the intensity and speed which you guys work out clearly seen from ChatGPT and all your breakthroughs and also in when you were temporarily removed from the company by the board and how all of your employees tweeted: OpenAI has nothing without its people. What would you say is the reason behind this? Is it the Biden mission to achieve AGI or something even deeper? What is pushing the culture every day? - I think it is the shared mission. 00:39:24,240 I mean I think people like, like each other and we feel like we've, you know, we're in the trenches together doing this really hard thing.

But I think it really is like deep sense of purpose and loyalty to the mission. And when you can create that, I think it is like the strongest force for success at least that I've seen among startups. And, you know, we try to like select for that in people we hire but even people who come in not really believing that AGI is gonna be such a big deal and that getting it right is so important tend to believe it after the first three months or whatever. And so that's like, that's a very powerful cultural force that we have. Man Thanks. 00:40:06,690 Woman Currently there are a lot of concerns 00:40:09,210 about the misuse of AI in the immediate term with issues like global conflicts and the election coming up. What do you think can be done by the industry, governments and honestly people like us in the immediate term, especially with very strong open source models? - One thing that I think is important 00:40:28,140 is not to pretend like this technology or any other technology is all good. I believe that AI will be very net good, tremendously net good. But I think like with any other tool it'll be misused. Like you can do great things with a hammer and you can like kill people with a hammer.

I don't think that absolves us or you all or society from trying to mitigate the bad as much as we can and maximize the good. But I do think it's important to realize that with any sufficiently powerful tool, you do put power in the hands of tool users or you make some decisions that constrain what people in society can do. I think we have a voice in that. I think you all have a voice in that. I think the governments and our elected representatives and democratic processes have the loudest voice in that. But we're not gonna get this perfectly right. Like we society are not gonna get this perfectly right and a tight feedback loop I think is the best way to get it closest to right. And the way that that balance gets negotiated of safety versus freedom and autonomy, I think it's like worth studying that with previous technologies and we'll do the best we can here, we society will do the best we can here. - Actually I've gotta cut it. 00:41:57,780 Sorry.

I know I just wanna be very sensitive to time. I know the interest far exceeds the time and the love for Sam. Sam I know it is your birthday. I don't know if you can indulge us 'cause I know there's a lot of love for you so I wonder if we can all just sing happy birthday. - No, no, no, no. 00:42:09,893 Please. - No? 00:42:10,726 (audience laughing) We wanna make you very uncomfortable. - I'd rather do one more question. 00:42:14,240 I'd much rather do one more question. This is less interesting to you.

Thank you. - We can, 00:42:20,150 you can do one more question quickly. ♪ Day dear Sam ♪ ♪ happy birthday to you ♪ - Thank you. 00:42:31,830 Somebody who's got a real burner and we only have 30 seconds so make it short. Man Hi. 00:42:41,160 I wanted to ask if the prospect of making something smarter than any human could possibly be scares you. - It of course does. 00:42:48,840 And I think it would be like really weird and a bad sign if it didn't scare me. Humans have gotten dramatically smarter and more capable over time. You are dramatically more capable than your great-great grandparents.

And there's almost no biological drift over that period. Like sure you eat a little bit better and you got better healthcare. Maybe you eat worse. I don't know. But that's not the main reason, you are more capable. You are more capable because the infrastructure of society is way smarter and way more capable than any human. And through that it made you, society, the people that came before you, made you the internet, the iPhone, a huge amount of knowledge available at your fingertips. And you can do things that you predecessors would find absolutely breathtaking. Society is far smarter than you now. Society is an AGI as far as you can tell.

And the way that that happened was not any individual's brain, but the space between all of us that scaffolding that we build up and contribute to brick by brick, step by step. And then we use to go to far greater heights for the people that come after us. Things that are smarter than us will contribute to that same scaffolding. You will have, your children will have tools available that you didn't and that scaffolding will have gotten built up to greater heights. And that's always a little bit scary, but I think it's like way more good than bad and people will do better things and solve more problems and the people of the future will be able to use these new tools and the new scaffolding that these new tools contribute to. If you think about a world that has AI making a bunch of scientific discovery, what happens to that scientific progress is it just gets added to the scaffolding and then your kids can do new things with it or you in 10 years can do new things with it. But the way it's gonna feel to people I think is not that there is this like much smarter entity because we're much smarter in some sense than the

great-great-great grandparents are more capable at least. But that any individual person can just do more - On that, we're gonna end it. 00:45:29,958 So let's give Sam a round applause. (audience applauding) (upbeat music)..