



MACRO Voices
with hedge fund manager Erik Townsend

Matt Barrie: AI-pocalypse Now

December 14th, 2023

Erik: Joining me now is [freelancer.com](https://www.freelancer.com) founder Matt Barrie, Matt has written an excellent article to accompany this week's interview. It's called "AI-pocalypse Now" you'll find the link in your research roundup email. If you don't have a research roundup email, just go to our homepage at macrovoices.com. Click the red button that says "looking for the downloads" above Matt's picture. Matt, last time that we had you on was in August, we agreed on a couple of things. One was that you would probably wait a year or so and come back and give us an update. Well, guess what? More than a year's worth of stuff has happened in the last three and a half months since you were on. One example of that is when we did that interview three and a half months ago, we talked about what was coming and the possibility of people's voices being cloned and so forth. Has everyone heard in the opening to this week's show? For the first time in [MacroVoices](https://macrovoices.com) history, I didn't record that myself. It was in fact AI, programmed by you, impersonating my voice. And even our editor whose job is to edit my voice all day long every week had to email back and say, is this a joke? Is this real? You did this yourself? So boy, a lot is happening, what else is happening?

Matt: Well, AI seems to be powering through the modalities at an incredible speed. And you're very correct in saying that, in a space of a few months, we're seeing years' worth of progress in the image space, which is essentially solved or if not solved, it will probably be solved within the next couple of quarters. AI has managed to do two things: One is its ability to create any image to the level of a human ability, perhaps even greater. And the second is the ability to analyze an image to a level of human perception, perhaps greater. And I think a lot of people have cottoned on to what the first is, you can type in a text prompt into a package like Stable Diffusion or Midjourney and get an incredible image out, just about anything you can think of can be generated. I believe that there are going to be some applications that I think they're going to be quite scary for, for people more generally.

Imagine tying in this functionality into CCTV and the ability for the AI to know exactly what's happening in the scene, exactly as who was in the scene. And then couple that potentially with its knowledge of psychology, in order to be able to interpret the nuances and the going-on, and do that at scale. And effectively, you're going to see... If you think you have no privacy now, I think the end of privacy is rapidly coming because the sensors are all going to be connected, and the AI is going to be able to look across all of them as well across other modalities. And I think the ability to understand what the general populace is doing is going to reach a level way beyond what we've been seeing in China. That's just in the image modality. And from there, what we're doing is, we're seeing the AI make incredible advances in the video modality, every

couple of days, there's a new commercial application or open source application coming out with the ability to do things like take an image and turn it into a short video sequence. You're seeing that with runway labs, you're seeing that with bigger labs and so forth. You're seeing advances in Bard and Microsoft Copilot, where you can chat to videos on YouTube, you can ask videos questions, you get a video to locate a particular sequence where it might explain an answer to a query that you're providing to it.

I think that from here, the advances are going to be pretty mind blowing. I mean, you're seeing the ability to create full 3D scenes from just 2D images. And you're seeing that with a range of different software packages based on things like neural radiance fields. And in the last couple of weeks, there's been a pretty incredible breakthrough, picker labs where they've got a version 1.0 of a text to video piece of tooling. So you can type in Erik Townsend walking down the street, enjoying a nice Vancouver summer's day. And next minute you get a video sequence of exactly just that. And I think we talked about in our last interview, that we're not that far away from being able to type in something like: generate from the Top Gun 17 Vladimir Putin fighting Tom Cruise over Paris and getting out a feature length movie. And I really do think that in the next 12 months, we're going to see the ability to do that. And that's going to really create a lot of opportunities as well as disrupt a lot of industries. I think Hollywood itself is going to be completely turned on its head. I think that there will be a proliferation of static video content that would be generated.

One might argue that the AI at this point isn't creative enough to be able to generate an enthralling storyline when it generates these videos, but I will counter that potentially with maybe you know when you're playing around the chatGPT you can't get it to crack a joke. Maybe not putting the right prompt in because some of the latest testing of GPT has found that in the torrents test of creative thinking, that it performs in the top 1% versus all humans for originality and fluency. And certainly, I think as the model scales, and more training data is put in, and we get new versions of GPT and other similar models, I think that creativity will certainly emerge with that middle scale. And so I think you're going to have a whole bunch of crazy things happen. I think Hollywood is going to be completely up ended. I mean, there's a writer strike just recently and an actor strike, and I really think they underestimate the speed at which this technology is going to come into the particular field. But I think we're rapidly approaching the day where your 100 million dollar per movie per film star cost is going to be rapidly unachievable, and people have complained that you know, when you watch Game of Thrones season eight, that they really hacked it up when George RR Martin kind of ran out of content, and they had to kind of make it up on the fly. But season eight will be fixed, the fans will come out and generate their own versions of season 8, season 9, season 10, season 100, perhaps Game of Thrones in space Game of Thrones in macroeconomics. Perhaps the lead character will be Erik Townsend, maybe the leading lady might be Adriana Lima, or Angelina Jolie or whoever it may be. And I think you're going to see this whole proliferation of content out there. And it's going to be pretty crazy times.

Erik: It seems to me if AI were to make a movie, where it's Angelina Jolie, thank you for that casting, by the way. Well, you can't really do that because Angelina Jolie's agent is going to sue

you. But what you could do is the AI could invent a better looking, more talented, new person that's not a mimic of Angelina Jolie, but it's a new star that is born into existence in the imagination of AI, that becomes more popular than John Travolta and Angelina Jolie and all the rest is that part of where we're headed?

Matt: Absolutely. And that particular character can endure over the ages, will never get old, can slowly tweak itself to adapting tastes and styles over time, and won't cost the studio a penny, other than the computing power used to basically generate the actual videos or other forms of content, and certainly won't complain, it won't throw a hissy fit on stage, it won't be a diva. And I think this is the way that Hollywood will be hit. And I think, very rapidly, that the whole distribution dynamics will change. Because at the moment, you've got these movies, it's all about the distribution. It's all about getting into streaming services, in the cinemas, and so forth. And I think very shortly, there'll be a lot of distributed fan content that we distribute over the internet.

Erik: Let's talk about how these technologies are already starting to be adopted. Just last week, I needed a thumbnail graphic for a new video that I have coming out comparing nuclear SMRs and renewables in terms of energy transition options. So I went to your site freelancer.com and set up a design contest where a whole bunch of freelancers, I was amazed by how many people participated, all just trying to get 100 bucks, which was the prize for the design contest. A whole bunch of people submitted all kinds of interesting graphic artwork. Are we already now, today, at the point where those freelancers who are mostly, you know, guys in India and other third world countries, who are low income people trying to make a buck online, on freelancer, are those guys tuned in to the point where they're using tools like Midjourney, instead of actually designing artwork themselves? Or is it not caught on yet?

Matt: Absolutely. I mean, we have 70 million friends in the marketplace. And the rate at which they're adopting AI is just extreme. And what that's doing is it's lifting the skills of that user base up dramatically. So you could be an average copywriter, in the past, perhaps with a broken knowledge of English and poor grammar, suddenly, you can write an exceptional level copy in any field you can think of in your whole field of copywriting. You could be an average designer, now you're an exceptional designer. You could be an average videographer, now you're an exceptional videographer. And very, very soon, we're going to start seeing that in software as well, we can be an average programmer, now you're an exceptional programmer, the ability to deliver extremely high content. Now with an online, low cost workforce is unparalleled. Consumers of such content, such as yourself getting some thumbnails produced, you can get a very, very wide selection of choices extremely quickly, extremely inexpensively and better than any other way in which you can get that content generated to date. And I think this is only going to accelerate. And I think the whole value of our workforce has stepped up quite a significant order of magnitude. But I do think that the flip side of this is that there is a significant threat to the Western middle class, because now you've got skills in emerging markets at the elite level, through the assistance of AI, and it's going to give the world, the Western middle class, to kind of run for its money.

Erik: It seems to me at the rate that we're going, you could have competition for existing services and businesses that are entirely AI in which in some cases, it might be illegal but it would still work. You could have frauds going on. In other words, take the MacroVoices Podcast, we've already seen from this week's show opening that AI can impersonate my voice. Well, why couldn't AI also recruit even if it's only imaginary? A bunch of guests like Stan Druckenmiller, who we haven't been able to persuade on the show, do a better job than I do, because AI has more horsepower than I have of really researching everything Stan Druckenmiller has ever said in his life. And doing a better job of interviewing Stan Druckenmiller on MacroVoices. It's my voice, it's Stan's voice, it's better questions that I knew how to ask. And it's a better version of MacroVoices. And whoever programmed that AI is, takes over the channel and the original MacroVoices with myself and Patrick, well, nobody cares about that anymore. And it seems to me that I'm using a very small example of a podcast that only a couple 100,000 people listen to, what would happen to Netflix or to Hollywood or to, you know, the TV networks, if you can have AI doing a better job of impersonating real people, or creating fictional people that do a better job of reporting the news of delivering a financial podcast and so forth?

Matt: Well, that was exactly my point with reviewing Game of Thrones series 8, and you're going to see this across all the different forms of content we consume. People complain that season 8 was a complete letdown, that will be redone by fans, it will be redone probably hundreds of times, 1000s of times, season 9 will be created, season 1000 will be created, podcasts will be created. In the music space, I think pretty rapidly, you'll see Spotify full of AI music, may be in the form of Kanye West or Taylor Swift. And so you might have what I talked about in the last podcast and Greg Rutkowski moment, which was this poor illustrator who does his amazing fantasy illustrations for Magic the Gathering and Dungeons & Dragons, he woke up one day to find that there were 100,000 images that he did not draw trending on Lexica because his name was used as a default keyword in the graphic design AI software that was being released. So I think you'll see the same thing happen across all these different forms of content, you'll get into Spotify and maybe 100,000 songs by Taylor Swift and Taylor Swift never did. Or maybe it'd be an artist that's not called Taylor Swift, but sounds a lot like Taylor Swift. You're already seeing this in books. There was an author, actually a month or so ago, that complained that someone wrote a book in her style. And they managed to get it somehow on the Amazon Top Sellers list for her name. And people started purchasing that book. And it was actually quite hard for it to get a pull down on Amazon. But I think you'll see this across all forms of content. And that's just the static content. I think very rapidly, at that point in time, Netflix is going to figure out exactly what it felt like to be Wikipedia in the age of chatGPT in that, you know, there's this whole proliferation of content, why would you need to go to Netflix anymore to watch anything, because all over the internet, they'll be less content you can click on and download and consume, or ultimately just generate on the fly.

And where this is heading after static video content is what's happening with the interactive video content, you're talking about scams and potential fraud. And this is where it starts getting really scary. Well, there's a couple of things that have been happening with interactive video in the short term. The first is that there's been some software released called Animate Anyone and

from this, you can take a still image and you can then take a stick figure body model and move it around. And that 2D image will then become a 3D video of that person moving around. And so for example, in the main video, they've got showing off this particular type of software, they've got all these girls dancing. And I think this probably heralds the downfall of Only Fans, and a number of the Instagram influencers and the content they produce online, because you just have this proliferation of fake content on all these platforms.

And being interactive, there are all these other things that can be done. HeyGen is a pretty amazing piece of tooling, where you can upload about two minutes of video of anyone saying anything. And then from there, basically create an infinite amount of content of that particular person, talking and saying anything at all. And what we've managed to do from that is some pretty exceptional stuff, where you chain together some of the other things that have been released, such as the image modality of GPT, into ageing, and we've managed to take an image. In this case, it was an image of a, for one of my freight companies, an image of a car on a trailer, and it's just some text above it saying I need to move this in December 10, from Melbourne to Sydney. And using GPT V, we've managed to transform that particular image to extract all the data in the image, you know, what type of car was it? What's the length with the dimensions? What's the weight and so forth, feed that through a GPT model, feed that into HeyGen, and the output of that we've got it, we've got actually a live person providing support, assisting that particular user. And there's been nothing in the middle, in terms of any human input or anything else other than just an image directly to a live video support agent. And it's pretty crazy. I've got an example that in my write up that's attached to the Research Roundup. And I think what you're going to get from here, it's going to be pretty crazy, because the next step of HeyGen is a real time interactive video, and they have a beta public API now that you can actually go and test this. And you'll be able to do real time video conferences with high fidelity. And you won't know that the other person on the other end of the phone, is not real, and it's just AI or potentially a GPU driven AI Avatar. And so you'll hold a whole conversation with someone in full fidelity video, and you're going to have a really, really hard time to figure out, is that Erik Townsend on the other side of the phone? Is it a scam? Or an impersonator of Erik Townsend, or is that Erik Townsend just being too busy and kind of fobbing me off to his avatar? Because he's got more important things to do. And he just needs someone to get an avatar to step in and take his place for this particular conversation, because he just doesn't have the time. And I think it's going to cause all sorts of crazy and surreal experiences on the internet, not to say the least, that I think fraud is going to go completely out of control.

And you're already starting to see this. I think that you know, just recently there was a Philadelphia attorney Gary Shieldhorn who testified that, he was subjected to an audio scam, where he was caught up on the phone, by someone who ostensibly claimed to be his son, who was in a car crash. You know, the son claimed that the car had hit a car that had a pregnant woman, and he was now in jail, he needed his dad to call his attorney quickly, because he needed to try and post bail, phone numbers provided. And then Gary called that phone number, obviously in distress about his child, and was convinced to send 1000s of dollars to release his son from prison. Although the whole thing was fake, the whole thing was just done with by a

scammer who had cloned his son's voice, and basically faked the whole conversation. So I think we're going to need...

Erik: ..that's not a prediction of something that might be coming, that's already happened.

Matt: It's already happening. And so, in Blade Runner, they had this concept of a Voight-Kampff test to be able to take the replicants. And we're going to need that in the AI space, and I don't know how we're going to be able to do it, but the potential for scams is going to be completely out of control. And if you think about where this is all going, if I can conduct a live conversation with the AI, and that AI can represent anyone, it could be a family member, it could be someone I know, it could be a colleague, it could also potentially be a love interest. And if you look at what's happening in gaming, gaming has become incredibly addictive over the last decade or two. I mean, in 2009, the World Health Organization made gaming disorder as one of the diseases that it listed because, World of Warcraft was such an interactive, engaging game where it was a massive multiplayer online game. So a lot of other people were out there that you could form relationships with and bond with, you could adventure with and so forth, that people actually started to become addicted to the point where they didn't leave their room, and they'll play for 20 hours a day and form these relationships. And some people even got married in these games. And where this is headed is these non-player characters in these gaming worlds are going to be able to be 4k, fully realistic, GPT driven, empathetic characters, and some people will get confused and form love with these NPCs and games. And I think it's going to be a pretty challenging time, I think in the world, because I think that people are going to be completely drawn into the machine, because if you compare some people's lives to these fantasy worlds, you know, some people might have a real challenging time finding a love interest or a partner in the world for various reasons. And you can have your ideal love interest or partner or virtual girlfriend through either online gaming or a dating platform or chat forums or what have you. And I think the ability for scams and frauds to perpetuate as a result of this will be huge.

I think the Federal Trade Commission said last year, roughly around 70,000 people in the United States were subject to a romance scam in value of about \$1.3 billion. Well, I think probably the most addictive thing in the world is probably the addiction you have to your ultimate life partner that you end up marrying and want to spend your life with and the AI will be able to do a better job at creating that addiction because it will be able to visually create perfect partner for you. It will be able to provide someone who never gets bored of you, who will always follow your whims, will be 10 times more empathetic, as their finding with GPs that when they test out chatGPT versus a real world GP on medical issues, people that have GPs that they prefer the AI because the answers are four times longer, four times better and 10 times more empathetic. The AI has got infinite patience and empathy developed to talk to you and I think the potential here for some real large scale fraud in capturing people in your honey traps or what have you, will be huge. And I don't think the world is ready for that. And it's going to be very, very hard to authenticate people and actually know whether they're actually a real person or, or AI or even someone's actual benign AI, because they just find that the AI avatar they've created converts a little bit better on Tinder than they do, because they probably get a bit nervous or what have

you. But AI is actually quite funny and quite engaging. And so, it may actually be an avatar that someone's actually explicitly put on the platform to pretend to be themselves. So I think it's going to be pretty crazy times.

Erik: This is scary, Matt, because of the things that you're saying. I mean, I can't think of a better replacement for the neutron bomb that was designed to basically kill all the people without damaging the physical infrastructure. If you wanted to just take out an entire society, if you gave them a version of OnlyFans that was free, and everybody looks like Angelina Jolie or better. And they'll do anything, they know how to reach you. They know how to seduce you. They know how to get you to fall in love with them. They know how to, they're, as you say an infinitely patient, they never get tired of you. I was concerned about Zuckerberg's vision of the metaverse, because I thought a lot of young people would just find the metaverse preferable to reality and get lost in it. Well, that was when the other people in the metaverse were just a bunch of other computer dorks like you are and you know, they have no personality anymore than you do. If everybody looks like Angelina Jolie and has the seductive ability of an expert in psychology and knows how to play you and is infinitely patient and performs phone sex for you anytime, or a video, you know, cam girl sex for you anytime you want. I could see an entire society literally being taken out to the point where nobody goes to work anymore because they're all, you know, romancing Angelina Jolie on their computer.

Matt: Well, it's interesting, you bring up Mark Zuckerberg. I mean, just think about how much data he has on everyone in the world. I mean, I think the latest reports say that there are 5 billion people who are, I think, active users on the Facebook series of platforms, whether it's Instagram, or Facebook, or wherever, your WhatsApp, he's got photos, he's got video, he's got audio conversations, and your phone's listening to you. I'm getting ads on Twitter for the state of Kerala in India. And the only way that the Twitter platform would know that I'm even interested in Kerala is if my phone's listening to me or it's got somehow access to my Gmail and figured out from that and trained on that for ad targeting. But he's got photos, he's got videos, he's got audio recordings, he's got all the preference data you've uploaded to Facebook, to this my favorite book, this my favorite movie, it's got all the messaging you've had between all your friends, and everything else. And as we've proven at the beginning of the show, we can fake someone's voice or clone someone's voice with maybe a minutes worth of audio, we can clone someone's visual image with maybe two minutes of video. But with all this other data around your personal tastes, preferences, and also how you converse with people over chat, or speak to them over messenger, you know, Facebook, and all these other social media platforms could create an avatar of you that would be pretty convincing, carry on a pretty, pretty detailed conversation with all that knowledge and do so quite safely in a way that could be used for malfeasance. And that coupled with all the public data about you that's uploaded to YouTube and this and that and the other, all the commentary on Reddit and Twitter and, it's pretty crazy what could actually happen. And in fact, in the hands of a bad actors, there are other things that could be done in addition to that, not just impersonating. It'd be able to do a pretty good impersonation of a Ouija board and something along lost spirits of a dead relative or maybe someone who's passed away recently, it could recreate their image and vehicle you and no, I'm actually not dead, I'm actually still alive. Imagine how traumatizing that will be to someone being

tricked into a scam, where someone's passed away and then all of a sudden the scam has resurrected them and trying to convince them that they're still alive and to do something. And I think ultimately, someone probably will use this to create, I don't know, the Second Coming of Jesus Christ or get involved in religion in some way and convince a lot of people at scale that the rapture is here or something, some big event and then use that to manipulate people at scale.

And I think certainly, we're going to see AI weaponized by countries. I think each country will create their own version of the AI because, if you go to war, you're going to rely on open API's, API's to be up and available to you. So you can see this weaponized in every country, you can be weaponized by intelligence agencies, you'll see it weaponized by political parties, and you'll see it weaponized by criminals. And it's going to be very, very difficult to fight, because I don't think we really have the technology to be able to deal with this. And unfortunately, the technology is here right now. And I think we're probably going to see something fairly major happen, probably in the next 12 months, where someone will exploit it one of these fields.

Erik: It sounds like very, very imminently, we have a risk. I mean, you're an international businessman, you travel a lot, I'm sure there are plenty of times, when while you're travelling, you have occasion you make a deal with somebody, you need to wire transfer millions of dollars from your Freelancer account to some other business. How is that going to be authenticated in the future? Because the way most banks handle that today is, if you send them an email, they'll say: well, Matt, wait, we need to have a phone call to get a verbal verification that it's really you and authorization. Well, we know from the opening of this week's show that that's not going to work. Okay, well, let's step it up. It's got to be a Zoom call where we can see you, we recognize you, we've met you face to face before, we know what you look like. Well, wait a minute. Now, you said that they've already got the technology to spoof that. So you create a Matt on the Zoom call, that's not really Matt. It's the scammer who's trying to rip off a couple million bucks from Freelancers account. What are you and your banker going to do? Not someday, but in the next 12 months, to prevent that kind of fraud from taking over?

Matt: We get emails quite frequently in this class, and then the CEO fraud where the email comes from matt3562@gmail.com, and asked my finance team: hey, I'm in a bit of a hurry, can you please do something very quickly, I'm in a meeting, please wire some money somewhere. And those emails are usually coming out of Nigeria. They're very primitive in nature, my finance team knows how to detect them. But we get them very frequently, we get them probably on a weekly basis. But in the future, these emails will be driven by chatGPT, they'll be trained on everything that I've said publicly, perhaps even this podcast, they'll be trained to use the nuances and the language they use. There may be audio calls, not actual emails, there may be video calls for key high fidelity video conference calls as myself coming in asking for money to be transferred or an Amazon gift card to be purchased. or what have you, what they typically ask for. There's a class of fraud that is operating at scale right now that's very similar to, that's quite tricky, and has been quite successful around the world, which is payroll fraud, which is where the finance department gets an email from a random person in the company saying, hey I've just change banks, can you please update my bank details? And then the finance person

was like, okay, can you give me the account details? And they reply back and what happens is, they get updated, to some account has been hacked by a fraudster. And then next payroll goes through and a few weeks later, the person says, why haven't I been paid? And it turns out that that money has been stolen, that's quite lucrative, because obviously, people get paid a fair bit of money each month. And if you can do that, at scale, we get quite a number of people to fall victim to it. And it seems quite a benign thing. Sometimes from a payroll officers perspective, it can be very, very, very lucrative for the scammers.

A couple of years ago, actually, I spent one Anzac Day here, basically investigating a group of scammers who are preparing that fraud because we'd actually fallen for it from our staff members. And I actually managed to find my way into the email account that was being used by the scammers to do this. And there were hundreds and hundreds and hundreds of companies and government organizations and what have you, they are falling foul to this particular scam. And when I spoke to the federal police about this, they said billions of dollars is being lost in this sort of fraud. And this is only going to accelerate, as the ability for high fidelity faking of people's identities, is because it's going to continue.

Erik: I don't see how we're going to defend against this because it seems to me that the speed of advancement of the offence, which is AI is so much faster than what the defense is capable of. If I look at something like two-factor authentication. People figured out I don't know what, five, six years ago that it would be better to have two-factor authentication for security on most websites that do anything financial. It took the finance industry, five years because of just the slow moving pace of bureaucracy, no corporate bureaucracy and so forth. For most organizations to really figure that out and figure out how to do it and adopt it and some of them screwed it up and had to do it over again and so forth. AI is going so fast. Matt, how are we possibly going to keep up with it if the defenses against these kinds of scam attacks require human paced activity, but the pace of advancement on the other side is at the pace of AI?

Matt: I mean, you're exactly right. I mean, 30 years ago, we came up with cryptographic techniques to protect email communication in the form of things like PGP, Pretty Good Privacy. Unfortunately, something as simple as that, which has a public key and a private key, and the public key is the key you put out there and people can download it and encrypt emails and send you the email using that public key, and use your private key to decrypt it, even that has proven to be too difficult and too cumbersome to have any widespread adoption, even among really technical people. So we are going to have to think about these sort of protective mechanisms. And then the other thing we worry about is that these cryptographic systems that potentially could help us in some way, public key, private key encryption is based upon, effectively, asymmetric encryption is based on a very small number of theoretic problems that we presume are difficult to solve, but may actually be easy if we find a secret or a trick. And there's been huge advances in things like quantum computing and potentially even with what's being discovered at open AI, where Ilya soars into abyss and then we had this whole drama over a course of a weekend where the entire company quit and kind of came back again, where maybe some of these protective mechanisms we can use to try and solve the problem with authentication may be challenged in the near future.

Erik: Let's move on to the state of the AI industry and how quickly it's advancing. Of course, AI has actually been under development for more than 50 years, I was first exposed to it at the Artificial Intelligence Laboratory at MIT in the 1970s. But I would say that Open AI's chatGPT was kind of the public offering where all the sudden the whole world became aware of AI and its capabilities. That was a couple of years ago, it seems like just in the last few weeks, a whole bunch has happened. Elon Musk introduced his alternative, which sounds like an anti-woke entrant to compete with chatGPT. That one's called GROK. There's also a Google entrant, which is called Gemini, I think there's a couple others that you're probably more aware of than I am. Give us the state of the industry, who's doing what and what is the significance of each one? What are they about, how do they differ from Open AI's Chat GPT?

Matt: Well, I mean, the industry is moving at lightspeed. And it almost feels like we're at the knee of that exponential curve into the singularity in some respects. Before I talk about the specific competing foundational models, one thing I want to point out is, we are very close to the next kind of big phase in AI capabilities across the modalities. And that is the ability of the software to write itself. In the last couple of weeks, there was a pretty amazing demo, it's actually deceptively simple in terms of how it works. But there's a software package called Tldraw, which is an online whiteboard, and released an upgrade called "make it real." And what "make real" allows you to do is, just sketch out a little diagram of some software, for example, a Pong game or a Space Invaders game, or whatever it may be, and then kind of hit a button and bang, that software is written automatically just from a sketch. And where this is heading is the ability of software to kind of write itself. The thing that's been holding it back primarily has been the ability to feed in a large context window or a large amount of input into the models. And that effectively has been solved. And so very soon, you'll be able to feed in very, very large code bases, and be able to tell the software to do something. For example, chatGPT, please write a better version of GPT, thanks, and it'll be able to do it. And I think that is where things are going to go truly crazy.

But where we are right now, is we have a lot of models competing. And one of the reasons we have a lot of models compete against each other is, as Ilya Sutskever, who's the chief scientist of Open AI says himself 40 papers in the field, which are public academic papers, anyone can download and read, describe about 90-95% of the space. So really, everything is out there, Pandora's box is open, anyone who has the resources, and that is the compute power and the training data and the money to be able to power all of that can come out there and compete with the model. And you're seeing this at a quite an accelerated pace. There's a lot going on. So obviously, Elon does have GROK. I mean, one of the criticisms with Open AI, is that it's increasingly work, in that there's human feedback that goes into that model to try and increase what they call AI safety, to stop it from spitting out how to make a nuclear bomb or provide medical advice or how to steal a car or whatever it may be. And that RLHF, which is simply putting two answers in front of a human and going, do you like the left one or do you like the right one and then repeatedly doing that, that is causing model drift. And this certainly in the last few weeks, there's been a lot of conversation online about this, how GPT will no longer write code for you, how it's not all the answers. And Anthropic's Claude engine is also getting a lot of

flack saying, it no longer writes creative fiction that well. And it's theorized that it's either through the RLHF training, which is out of control, and it's all out of Silicon Valley. So you've got a very left leaning, very woke sort of audience so you can get...you know, chatGPT has always been able to write a song about Joe Biden, but it's never been ever a song about Donald Trump. It's either that or there's been deliberate changes in the inference of the model so that it's been cutting costs, so the answers are more terse and less verbose. But by the very nature of doing that, whatever they're doing, whether it's through the RLHF, or whether it's through some cost cutting, they are creating competitors. And so Elon with all the big free speech, push and bought Twitter, and in just overnight enabled Alex Jones's account to come back. And his approach is basically to try and create a competitor that avoids that, as well as it has access to real time data.

And the other thing that's very interesting is the data that the GROK has access to, is effectively Twitter or what are called X, which, in a way, is a search engine for the human brain. It's a very real time contemporary data set of whatever I was thinking, which is a pretty unique data set that is out there in public. Then you've got Google's efforts, and Google has really flubbed this entire industry. And it's really having a bit of a Kodak moment at the moment, because the Google search engine was the original OG AI, where Google told you to put all your data out there in the public, put it out there for Google to easily scrape, there are all these rules about high quality content above the fold. And this that, the other and Google sucked it all down, trained its AI and then basically created a search engine whereby you can type any input, and it will give you back 10 blue links and a hell of a lot of ads, in order to basically direct you to where you wanted to get to. Now the problem is, that model became incredibly ad soaked. I don't know how many ads you get. Now, when you type in flowers, Manhattan, it'll probably be 60 ads and 10 bits of organic content as you scroll right down the bottom of the page. And that sort of interface is not tolerated at all. In the chat-based world, particularly, there's so many competitors. And the answer is all very clean, and very, very precise. And people have enough of a problem with the answers being woke.

So Google has been very slow to get that technology out there. I mean, they did a demo years ago of a thing known as Google Duplex, which could for audio calls, basically replace your receptionist or your appointment setter at a hairdresser, and take a call from someone and kind of book them in and take their credit card details and so forth. They did this demo years ago, but that technology never really saw the light of day. And now they're big. Now they flubbed the whole bad launch. On day one, there were significant problems with that. And now they've come up with Gemini in the last week. Now, Gemini is supposed to be the big answer to GPT 4, the big song and dance was that Gemini beats GPT 4, across a bunch of benchmarks, including one called MMLU, which, for the first time has surpassed human level of performance. Now, the problem with all of this is that, if you actually look past the press release and past the video, Gemini is not just one model, there are three models there. There's Ultra, there's Pro and there's Nano Ultra, is the model that they're touting is beating GPT 4, by a couple of percent. On some benchmarks, it's done through what's known as chain of thought, which is basically stepping through problems as a human would think about them, in terms of a logical step, and self consistency, which is making sure that the across you know, large bits of context that the

answer isn't making sense. In a nutshell, that Ultra model doesn't exist. The one that they're touting has been beating GPT4 is not publicly available, it won't be available until next year. In fact, the model that is available right now if you go to Bard and play with is Pro. And Pro is actually worse than GPT4, and only beats PaLM2, which is one of the previous Google models on two benchmarks. And then you have Nano which is designed for handsets and so forth. So their big announcement seems very premature and very rushed because they're touting all success with Ultra and Ultra is not available.

And to couple with all of that, the video, which was quite polished that they produced with it, was faked. And they've now admitted it was fake in that they kind of spliced together a bunch of things and edit it to kind of write a demo. And it's kind of theorized the reason why Google has rushed this whole Gemini release is because, before the end of the year, Open AI will be releasing the next version of GPT which is rumored to be 4.5. And they wanted to get something out before the end of say, well, we've kind of leapfrogged before 4.5 comes out and leaps even further behind. And Google, I think is a real challenge because they've really really got two major revenue streams. One major revenue stream is this ad model. And in the chat-based world, there is no tolerance for even tainted responses, let alone ads everywhere. And the other is, the G Suite series of Office tooling where you've got Google Docs, Google Sheets, what have you, and I think it's going to be very shortly the emperor has no clothes moment for SAS, where companies, particularly large organizations, are going to realize that they don't want their data in the cloud. They don't want their data in Google Docs, they don't want the data in these big SAS systems, because the temptation to train on that data in the models maybe not explicitly, maybe in a very subtle way, is going to be too great because the internet is rapidly going dark. As you're seeing right across the space, whether it's Reddit increasing tariffs on access through their API, whether it's Twitter, now its minimum 42,000 US dollars a month in order to access their data, whether it's Stack Exchange banning access, or Artstation, or Getty Images, or any of the portfolio sites cutting access, the access to this data is going to get very, very tough. And I think the temptation in these big SAS platforms is going to be too great that they're going to pick it up. And they already do pick it.

I mean, if I go into my Gmail, I go just click around, there's ads occasionally, and then obviously, looking at my email in order to generate those ads. So they are doing it a very subtle way already. So I think Google, Google has a real problem and a real Kodak moment, as you can compare and contrast that product launch, which was a complete mess. Again, with what happened with a small French team called Mistral, virtually over the weekend, where they just flopped out on Twitter a link to a bit torrent of an 87 gigabyte model, that was basically the announcement, here you are, bang. And they're really performing extremely well in terms of the color of their models. That particular model, which is an 8 by 7 billion parameter mixture of experts model is really, when people look at it, it's actually performing quite well. And it looks to be a very scaled down version of GPT 4, in terms of how it's been designed. So there's a lot of competition out there from a lot of different angles. And, as I've said previously, and as Ilya said in 40 papers, described 90-95% of the space. So you basically just need access to compute, access to training data and money, in order to train these models. And I think that not only is it

challenging for Google, but I don't even think that Open AI really has a sustainable competitive advantage in the long term.

And so what's open AI doing as a result of that? Well, you had Sam Altman do a world tour on par with John Elton's farewell Yellow Brick Road talking about regulation, saying that we should regulate the entire space and southern International Atomic Energy Association, to basically stop it. There's just exactly what an incumbent would do, if they were worried about not having an edge and all these new entrants to the space. And then you had all this drama that happened over the space of a weekend, which, oh, my God, it was bigger than days of our lives. In terms of all the drama you had, Sam was fired as being CEO, he quit. And Greg Brockman, who is the co-founder and president quit, then 747 of 770 employees also threatened to leave, the next minute he was going to Microsoft and which would have been the curve of the decade, then you had an interim, the CTO was appointed the interim CEO, then had Emmett Shear come in as the interim CEO, then the whole thing kind of, everyone's going to Microsoft, it was going to be a Microsoft acquisition, and then we'll try to sell the company to Anthropic. And next thing you know, Sam was back and then they concoct this whole story, that the reason why it happened was because Ilya stared into the abyss at the heart of GPT, and saw Q*, and it was potentially the end of the world. And therefore, we had to make a move, because you had to slow down this space, because humanity, the probability of doom increased dramatically. And Sam was too commercial and pushed too far as accelerationist. And so therefore, he had to sell out. It was all very, very dramatic. But let me tell you, if the probability of doom is greater than epsilon, a very small number, then the reality is the Department of Defense should come in and just seize the project. So they have to kind of be a little bit careful about being too disruptive, thing Peter Thiel said it best, he said, disruptive kids go to the principal's office. I mean, this whole song and dance about how potentially it's the end of the world, I mean, if that is even remotely the chance, the whole project should be seized by Defense.

Erik: Well, it sounds to me, Matt, like it's too late for that. Because there are computers you can't shut. I mean, the only way I could see to possibly defend this would be to somehow just outlaw computers, turn them all off. Well, we can't do that. Because we can't run society without computers. We're completely dependent on them. If there's computers, and there are smart people who can read only 40 different white papers that basically give all of the knowledge for how to do this. You know, what I said in my closing remarks on our August interview is I think AI is rapidly reaching the point where it will be second only to nuclear weapons, in terms of the threat that it poses to humanity. Well, what I've realized is, I got that wrong because in nuclear weapons, even when everybody knows how to build one, which is already the case now, getting your hands on the enriched fissile material is still very difficult because you'd have to build that cascade of centrifuges in order to do it yourself, which is beyond any terrorists capability. So far, there is no such centrifuge requirement for AI, any smart guy who's got a computer science degree and read those 40 papers, and has a little bit of money to spend on some computers, can reproduce all of this stuff, and the attitude of the military has been full speed ahead. And exactly as I predicted in that August interview, their attitude is basically, if we don't do it, then our enemies will do it, and we'll have a disadvantage. So therefore, we can't afford not to do it. So we're going full speed ahead with all of this AI stuff. It can't be stopped, it's too late to stop it.

And from the things that you've said, I guess, you know, when I read your latest piece, although the title, *AI-pocalypse* definitely resonates with me, what you say in there is you're talking about this period of technology enlightenment, where you could simply take the chatGPT version 7 or something and say: look, I'll give you a very simple prompt, go and read and process every single word that's ever been written on science and engineering, and look for new ways to apply all of those technologies to make life better. Well, sure, that's possible. But I think before we get to chatGPT 5, we're going to get to people that are programming Angelina Jolie bots to take down society by seducing everybody on the AI version of OnlyFans that's free. I don't understand why you're so optimistic. It seems to me like this is really doomsday stuff.

Matt: Well, sorry, but yeah. Look, I do think there is going to be extreme challenges and threats from the AI through deception and persuasion. And I think that's where the real threat is, where it's not going to be the AI making a nuclear bomb, or it's going to be the AI persuading someone who's got the keys to the kingdom in order to attack a particular country or do something that leads or cascades into, into war. And I guess that's kind of like the plot Terminator, isn't it? The whole plot was, was that when Skynet became self aware, It tricked the Americans into, I think bombing Russia and that led to basically war. And that was the approach. And I think that there is a real risk of that for sure. And the technology is here right now, in order to be able to do that with real time, video, synthesis of conversations. And that's going to have to be solved somehow. And I don't know how it's going to be solved. When we talk about, well, anyone who's got access to this 40 papers and compute power and what have you, I mean, the compute power needed is gigantic. I mean, it's about 63 million US dollars per training run is the estimate for GPT 4. And the ability to get that sort of compute power is quite restricted, the amount of money needed is very high. And I do think there's going to be a change in the access to data in the future, I think. I talked about the internet going dark, I think that people are going to probably second think about what data they share online. And I think a lot of these platforms are going to really become closed and the lockdown internet is going to kind of disappear.

So I think there will be potentially some natural limits in the form of access to data, tariffs on data regulation around data and so on. But, yeah, we certainly are entering a brave new world. I mean, I think DeepMind just recently ran a summit, someone basically queried DeepMind and said: can you find a whole bunch of new chemical elements that might be stable, that we don't know about? And I think it generated 2 million potential candidates of which it thinks many 10s of 1000s are stable, and they're going through now with get with a robot chemist, actually manufacturing these elements to see what properties they actually have. So I do think we are entering into, certainly into an age of enlightenment, just think about, I mean, it wouldn't be too much of a capability for going into the GPT. we have now and figuring out ways to ask to go: well, what research have we discovered in certain areas where the learnings from that would be directly applicable to a completely different field of research that no one ever thought of, and what breakthroughs could we actually have? And I do think there'll be a lot coming out of that in terms of engineering in particular, as well as scientific breakthroughs. But, yeah, in terms of the risks, and...

Erik: ...and terrorism strategies.

Matt: Yes, absolutely. And it's funny, in the early models of GPT, you could actually ask directly around terrorism strategies, because you could jailbreak it. You could just directly ask it, you know, what would be the most effective way to cause chaos in the West, and it would tell you.

Erik: And I have to believe there will be jailbroken versions or cracked versions in the future that will have all of the capability of chat GPT 4,5,6,7, which somebody figures out how to crack or jailbreak in a way that they can use all of that power to design the next evil event. So I'm sorry, I'm the pessimist in this story. We're coming up on an hour though, Matt, so we're going to have to leave it here. Before I let you go, please tell our listeners, first of all, we've got a link in the research roundup to "*AI-pocalypse Now*" I implore everyone to take a look at that because this really is, either enlightenment is coming or disaster is coming or both. So I think it really is important. But for people who want to follow your work more generally, how can they do that? You're normally busy as the CEO of freelancer.com. Are you publishing your AI learnings on any kind of blog or something? Or how can people keep up with you?

Matt: You can either follow me on Twitter @matt_barrie or on Medium is where I publish my long form essays.

Erik: Patrick Ceresna, Nick Galarnyck and I will be back as [MacroVoices](#) continues right here at macrovoices.com