**CRAIG:** Hi, I'm Craig Smith. And this is Eye on AI.

Before we begin, I want to shout out to our listeners beyond north America and ask that some of you get in touch. I'd love to know who you are and what you do.

In Africa, we have listeners in Ghana, Tanzania, Uganda, and Somalia among other countries. We have listeners in Iceland and Nepal, a handful of people tune in from Bangladesh, others from Kazakhstan. If you were listening from one of these countries, or any country outside the US, email me at [craig@craigsmith.ai](mailto:craig@craigsmith.ai). Maybe we can do an episode with listeners from around the globe.

This week I talked to the inimitable Ben Goertzel, who anyone paying attention to AI has seen at conferences or on YouTube with his long curls and cheetah-print slouch hat introducing Sophia, the AI enabled talking head, or waving his hands as he describes the path he believes will take us to human level artificial general intelligence, otherwise known as AGI , the holy grail of AI research. Ben is a curious character, one that I've been fascinated by since first hearing him speak in 2017. He oversees a global network of researchers working on a sometimes confusing array of disparate projects, most under the umbrella of his nonprofit foundation SingularityNET.

When I met him, he was living in a village house in a rural area of Hong Kong, and he now lives on an island off the coast of Washington state. He spoke to me about Sophia, his talking Philip K Dick avatar, his work in medical AI, his frustration over Big Tech's dominance of AI research and his own quest for AGI.

Get ready. Ben talks fast and covers a lot of ground. I hope you find the conversation as beguiling, as I did.

**CRAIG:** I'm very interested in the drive by many researchers into unsupervised learning, different forms of unsupervised learning.

I'd like to hear generally your thoughts on unsupervised or self-supervised learning, learning that doesn't depend on labeled data.

And then personally, I'd like to hear what's going on with Sophia, because I see you on the internet all the time with Sophia, but I've talked a lot to people in NLP and the chat bot world and I don't understand what's going on with Sophia because I know that chatbots are not that sophisticated.

Is Sofia truly connected to an AI that when you speak to it, it is speaking back.

**BEN:** She is, it is a topic that people are very confused about in, in many different directions, both thinking it's more than it is and less than it is.

Sofia, obviously it's David Hanson's creation more so than mine. Although I was, I led the software team at Hanson Robotics for a few years.

And the first thing to understand, that's a hardware platform and it's a mushy ware platform with a face. It's flexible and you can control that same robot with a host of different software systems. And if you're a naive observer watching a video with a robot, there's really no way to tell what's going on behind the scenes, in any given instance.

So, when Sophia is on a major talk show or something, it may well be that's entirely scripted in advance. Of course, the humans may be entirely scripted in advance also and reading a script, reading off a teleprompter. There could also sometimes be some AI driven chat in there, but by and large, in such encounters, there's a lot of scripting going on. And when Sophia is standing up, giving a speech to the UN or something there's a lot of just someone typed in a script and Sophia is reciting it. And that's not AI, right? It's very cool robotics. Now, we also a number of times, we ran a fairly sophisticated AI system using our OpenCog platform with Sophia as the input output and then whatever is said is going to some knowledge graph and there's some integration of visual perception with language. And, we ran that in some trials and published a paper in one of the NIPs conferences which was Sophia giving them like meditation guidance. And that was really quite cool because people got into some profound trance states after talking to the robot and that, that was running this whole neural symbolic AI platform behind Sophia.

In some ways more sophisticated than Alexa or Google Assistant, in some ways, probably less. But anyway, it's a real AI. Now much of the time, Sophia is running the Hanson AI chatbot system, which is intermediate and subtly between those two extremes, it's not OpenCog, but it's also not just scripted.

There's a rule based chatbot in there. They have their own chatbot system called Soultalk, which has a whole bunch of if-then rules with variables that are organized in a bunch of trees. There was also a transformer neural net in there that it falls back on sometimes. Right?

So, a GPT-3, or a BERT type thing that they train particularly. And then there's some neural models to deal with visual perception. So, it's really, it's like a messy hybrid system with some rule-based chat, more like a game tree dialogue system, but sometimes it will call on a generative neural model to utter something.

I know the rule-based pretty well, so I can sort of tell when it's the neural model versus the rule-based tree but most of most people couldn't tell. I don't think that system is especially less sophisticated than that a Google Assistant or Alexa actually, which has also has a bunch of rules and models mixed up.

And it's not trying to be as useful as those are, but it's trying to be more emotionally evocative than those are. And so, then the confusion's there because there's a beautiful human face associated with it. Some people are like, oh yeah, she really understood me. And on the other hand, some people are like no, that's all puppeteered by some human there's something, which is, it's not the case either.

The reality is it's another like hybrid's dialogue engine, which occasionally really does understand something it's saying. And often doesn't by my subjective interpretation understand what it's saying, but it is just giving some plausible sound. That's just, the reality is boring and complicated, so no one wants to pay attention to it.

And then they want to either say it's a fraud or it's an AGI, which it is not.

**CRAIG:** Yeah, I can see for Hanson robotics, it's been a wonderful marketing tool. What's the why? Why is it out there? Why…

**BEN:** The why is a bit complicated. David Hanson is an artist, right? And so, he started out wanting to just bring his sculptures to life.

So, in a sense the original why was really the same why have an art installation or a sculpture? That's really what drove David as a person to do it. And then he has a whole singularitarian ethos where he believes the singularity is going to come. He thinks AGI is going to be smarter than people.

He wants them to be loving and compassionate and as an artist, his intuition is you want the AI's and robots to empathize with people and that's going to lead us in a beneficial direction. So that's ultimately the why is David Hanson as an artist and visionary. And then when you wrap it into a business, there's a PR goal and obviously there's a goal of moving towards scaleably manufactured like home and office service robots, which will then lean in a little bit different direction than Sofia is.

So yeah what we're doing with Hanson robotics now is a joint venture between Hanson robotics and Singularity Studio, which is one of my companies I'll tell you more about in a few minutes. It's a joint venture aimed at basically humanoids nursing assistant robots for eldercare.

So that's, it's the same essential hardware platform as Sophia, we're going to sculpt the slightly different face for it. But the objective is different and both more and less ambitious cause this is aimed just to be like a useful robot to provide some companionship, some conversation, some advice, listening and help with scheduling things in the house and so forth.

That's an example of something you can do with the Sofia platform, which in some ways it's easier and some harder than making a general purpose chatbot, right? Because on the one end, it doesn't have to hold forth about sports or politics or anything. On the other hand, it actually needs to understand what's going on with that patient in a sense.

So that's the one I'm involved in because it interested me more because I’ve done a lot with biomedical AI. They can roll out like car salesperson robots, and so forth. You can view Sophia is a showcase and then it's a platform that can be used to make domain specific products. So, in terms of pure research, I found it very frustrating because Sophia has to pretend she knows what she's talking about when she doesn't. And if you're trying to do AGI research, you want the system to be talking only about what it actually understands. So, you can teach it. You don't want like a bullshit artist, right?

**CRAIG:** Yeah. And for real AGI research, the platform is irrelevant, right?

**BEN:** No, I don't know that the platform is irrelevant. There's an argument that having a robot that can interact whenever humans are in a shared physical environment is really useful for teaching an early stage AGI.

**BEN:** Actually, I think the platform is great. You could do something different with it, right? Like you want the AI to be learning about why this is a cup rather than a bowl. And like how much stuff will come out of it if I spill it over. So, you want the dialogue to be grounded in the robots experience rather than using the robot as an art piece and whatever random thing people want to talk to the robot about.

So, the use of robotic hardware and humanoid robotic hardware for AGI research is, it's arguable, but it's certainly not, it's not trivial or dismissible. It's interesting. But the, yeah, the use of a robot character that has to give the illusion of having an understanding it doesn't have for AGI research, that's more complicated and in some ways is a negative use, right?

**CRAIG:** Yeah. Actually, that's what I'm getting at because a lot of people know that I'm interested in machine learning and they're like, oh my God, have you seen this Sophia?

And it creates, in my mind, a lot more confusion than understanding.

**BEN:** Yeah. Although there's a lot of machine learning on the back end there, of course. There's computer vision for face recognition. There's reinforcement learning for movement, there's transformer neural nets, but on the whole, they're combined together in a particular way.

But again, that's true in Google Assistant, too. And that, that also confuses some people in a different way, actually. It confuses my two year old son as to why the thing is so smart yet stupid.

MUSIC

**CRAIG:** We'll get off this topic, but also the Phillip K Dick video.

**BEN:** That was like 90% transformer neural networks. That was OpenCog plus transformers. So, we made a very simple Philip K Dick chatbot in OpenCog, the system is wrapped conversational dianetic. Then almost everything that PKD said was from a transformer neural net we trained on Philip K. Dick's exegesis and his letters and various novels. So that was fun. That was the voice was trained from Philip K. Dick speeches, also. So, it was a separate, like audio transformer and seek to seek model. That was really an art object, right? David Hanson and I both love, we both love Philip K Dick, and I just thought, the weakness of modern transformers is they generate a lot of bullshit.

So, I thought if you're generating something that is supposed to be like a stoned philosopher, just saying weird, cool stuff, then that becomes a feature as much as a bug. And it's cool. And it did work out that way. Like it, it rambles all sorts of wacky out there stuff much like Philip K. Dick in some of his writing.

In the medical, actually we're battling against transformers in a different way. We have one interesting experiment I did like a few months ago. I just set up a huge online dialogue between Sophia's chatbot and the transformer dialogue model similar to Facebook's chatbot with some improvements, but just to see how the two chatbots chat back and forth with each other.

The neural model, on the whole seems a little more fluid in terms of the diversity and syntax that it throws at you. And then, it’s much worse than keeping the thread than Sophia's chatbot with the rule based aspects like digresses more. And of course, the transformer models have the feature that like 10 to 20% of what they say is complete and utter bullshit that makes no sense, but it looks like it does.

In this eldercare robot use case, it's very bad, right? Because you can't have your eldercare robot say even 1% complete random bullshit. So that actually that means either you need to use a rule based system instead for dialogue. We have some rules in there because there's some content, which is just, it has to be a certain thing given the nature of the domain of diagnostic questions, are there procedures of a hospital or something. But we don't want to get into encoding a huge amount of rules for a medical assistant. On the other hand, we also can't have it bloviating random things some of which could be inappropriate or dangerous even, right?

You need a way to use understanding of the world or structured knowledge to guide what the generative model is doing. And so that's one of the things we're working on in the context of this project called Awakening Health, this elder care robot project, and there's two approaches there, right?

One is you give a bunch of negative examples and you try to train the transformer model to avoid anything like that in certain states. The other, which is more interesting, you can take a tree transformer type approach and you give a knowledge graph, and then you use the knowledge graph to bias the generation of the generative model.

Like in a true transformer, they have a simple constituent grammar biasing generation, but we can have a whole OpenCog knowledge graph biasing generation. Like that applied project, it's not AGI, it does force us to do neural symbolic stuff and to push beyond the normal neuro models. But see, to do that with Phillip K Dick, it's too hard. What's in that knowledge graph, like the universe and everything, whereas with the medical robot, what's in the knowledge graph is, there's actual structured ground truth data about drugs and diseases and schedules and so forth that you could use to keep the generation on track.

**CRAIG:** Fei-Fei Li is doing lot of research on ambient AI with multiple sensors. And, you could combine that with other kinds of sensors and have the robot ask questions when it senses that something's going wrong.

**BEN:** In some of my more like corporate consulting stuff, I and my team built like phase three ID systems for big box retailers would recognize it's the same guy walking out of the scope for this camera, into the scope for that camera. So that same software you could use for the cameras in the hospital. We can feed into AI that, that the robot mind is here.

**CRAIG:** That'll be fascinating.

I have to tell you a little anecdote. My mother who's since passed away was in the hospital a few years ago, and the hospital rolled in this very crude looking robot. It had two camera eyes. A little speaker mouth, and it would sense movement. And then a nurse, rather than having to run down the hall would get on a microphone and say, are you okay?

Not my mother, but the person next door, the robots spoke up and the woman lying in bed said, ‘is that God?’ Which I thought was interesting. The understanding of what exactly is happening for an elderly patient.

**BEN:** We've been doing some investigation with elder care facilities and working with a company in US called Connected Living, which has IT infrastructure rolled out in hundreds of eldercare facilities.

There's a lot of people there who know all the options on their iPhone, better than you or me. They're not doing much else all day, and they're quite sophisticated. And then there's another segment of patients who forgot their kid's name making a robot to deal with both of those cases is a very interesting challenge actually.

**CRAIG:** Can you just for listeners describe OpenCog?

**BEN:** Sure. So OpenCog is a, AGI oriented software platform, which was founded in 2008, but actually some of the code existed, it just hadn't been opened up yet, in 2001 or something. And the notion is that AGI can be approached via cognitive synergy by a synergetic interaction of a number of different AI algorithms on a common dynamic knowledge store where the knowledge store is this weighted labeled hypergraph called the AtomSpace.

And then the different AI algorithms cooperating together on this AtomSpace, they have to be carefully designed, so they don't step on each other's toes, but work cooperatively. And when one of them gets stuck, the other ones help him out. And this transcends distinctions like neuro versus symbolic, because this knowledge hypergraph has some symbolic links that are morphic to say predicated or term logic instructions.

It also has neural net links. So, you can do neural and symbolic together in the same distributed knowledge hypergraph. We've been playing with this sort of thing since the late nineties, we've made a lot of progress in lots of different application areas. And I'm really curious to see what happens when you massively scale this up, which is what we're pursuing in the TrueAGI project.

SingularityNET helps with infrastructure for scaling that up, but there's a whole bunch of other interesting things that we need to pursue. It's a very different direction than what the mainstream of the AI field is taking now, which is mostly focused on like back propagation through huge, like multinodular neural networks.

So, we can connect neural nets learning through back prop with the OpenCog AtomSpace So, if you're doing probabilistic programming, you can have the same problematic program act over part of the AtomSpace and some deep neural nets, but it is a different way of representing and organizing knowledge.

**CRAIG:** That's fascinating.

MUSIC

Okay. Let's talk about where you're spending most of your time.

**BEN:** Let me give you the big picture. So, I've been leading as CEO, this project called SingularityNET since 2018, it was founded in 2017. That's a blockchain based AI infrastructure company, right? So, we released a platform that basically lets you make a decentralized society of minds with an economic aspect where there's, it's helping to pay AIS in the network for services.

Although you cannot pay in regular money and it's integrated with a token on the back end. And then the AIs can outsource work to each other on, on the back ends. So that platform exists. And then from a business view, the big challenge is getting traction for that platform. Like why is anyone going to use it?

And there you have the same problem of any other platform in the tech business, like an Airbnb or Uber or something. The more AIs in there, the more people will use the AIs, the more users, the more AIs will go in there.

So, what I've been working on, actually SingularityNET as an entity is spinning off a bunch of other companies, mostly for-profit companies and partly because we've been incubating these things internally inside SingularityNET, and they just need more resources, not just money, but more collaborators to grow and partly that's one way to get traction on the platform is like venture studio type of growth, like spinoff companies that will then will then use the platform.

And then hopefully that will kickstart the growth of it. So, these are all of course, AI related in various ways. And so, I've described one of these spinoffs, which is Awakening Health, which is doing these nursing assistant robots. That's a joint venture with Hanson Robotics. So that's I hadn't been doing that much with Sophia or Hanson robotics for a couple of years, actually, although we did this Philip K Dick thing just for good fun.

But now we're plunging back into Sophia-land with a bunch of new AI technology and then another, so another of these spinoffs is called Singularity Studio. And this is basically, it's like an Element AI or Palantir type thing. Palantir, but focused on doing good. It's enterprise AI, and we've been looking at the healthcare space.

So, we've been, we're working with some folks doing COVID clinical trials. They're doing trials of like cocktails of multiple antivirals. And then we use machine learning to figure out which antivirals to give someone based on gene expression data. And we've been looking at some breast cancer stuff in the same context. Basically, using neural symbolic methods or fancy neural methods like InfoGAN, but also some probabilistic reasoning to try to do transfer learning better from one clinical trial to another. Because the current state of things when you do ML on a clinical trial, it's very hard to get the models to transfer even to another trial of the same drug on a slightly different patient population. Let alone to do the transfer to a different drug or something. So, we're trying to get transfer learning to work for personalized medicine trial analytics.

And with interesting success actually like this is the first thing I've gotten InfoGAN to work in a useful way, which is that's a kind of generative associative network with unsupervised learning. Okay. An infoGAN automatically learns structured, latent variables as part of the generative model that tell you semantic features of the data it's modeling.

So, like an infoGAN for face recognition, it'll automatically learn one variable for how wide the eyes are, one variable for how long the nose, one variable for how fat the lips are or something. So, it's learning that unsupervised. What are the key features of the things it's generating? So that seemed really cool when it came out, but I never found anything was actually more useful than something that models for until clinical trial analytics, it actually seems to be. And I think the reason is there you have very skewed samples relative to the complexity of the data. In face recognition, it's cool to see that the system by unsupervised learning will recognize the key semantic features of the face. But it doesn't matter because you have too many faces relative to the complexity of the face. For a clinical trial, you have a couple of thousand patients, but the genome of each patient is large and there's so much diversity among people.

So yeah, what we're looking at there is using these infoGAN-type models to extract key features from patients as related to the therapies they're taking. And then we use sort of logic reasoning to try to come up with a causal explanation. Like what why are these semantic features there?

Why are they important for predicting if someone's going to be cured better by this drug than that drug whether they are a breast case or they are COVID. So that's not AGI either, but it's being very interesting and we're seeing a big value for neural symbolic combination there, which is interesting.

I think biology is good for that because the data is scant. And so, you can't do like you do with face recognition or GPT-3 or something. Cause you just don't, you just don't have, we don't have gene expression data for every human on the planet in our data at this moment.

Hopefully we will eventually.

**CRAIG:** Yeah. And where is this research being done?

**BEN:** This is being done with SingularityNET, SingularityNET Foundation, but then Singularity Studio is spun off as a for-profit company. We're working with the hospitals and pharma companies and such, just doing data analytics on a commercial basis.

Physically we're all over the place. We're in our biggest AI team is in St. Petersburg, Russia, but where are you in and Belo Horizonte, Brazil, in Ethiopia in Addis Ababa, in Bangalore in India and then Hong Kong. Few of us scattered around the U S but we don't have a central presence.

**CRAIG:** And so, this research, is there a central platform that you're all working on or is it just emailing and working through Slack? I'm just curious.

How does the coordination work? We use GitHub and Slack? Same as everybody else, really. In terms of the software, the SingularityNET platform, and there's OpenCog engine and then the whole world of neural nets. We have a server farm in St. Petersburg, which we've customized heavily to be relatively efficient for the kinds of AI that we want because using AWS and stuff for this is very expensive.

**CRAIG:** So, it is Kind of a private cloud.

**BEN:** Yeah. For transformer that's highly demanding on like multi GPU, but then OpenCog is highly demanding because you you're given these big graphs. And then of course AWS or Azure can do anything, but the cost per hour becomes very high.

Yeah. There's other things, we're launching an app on the App store, which is also a medical focus which is a to Rejuve app and that's from a different spinoff project called Rejuve. It does a bio signal analysis on data from your apple watch, like pulse oximetry, HRV, and also temperature from a digital thermometer.

And so, we do anomaly detection to detect infection before you feel sick, if there's something screwy in your vitals. So that's again pretty interesting. And not on my main AGI research direction, but it's clearly quite valuable. And that's, the more data you can collect there, the more interesting stuff you can discover right? Now there's no Bloomberg of health data, right? There's no huge repository of data that you can use to drive forward machine learning and machine reasoning in the, in these contexts. So yeah, SingularityNET is going into other stuff. We're launching a big decentralized finance project, which is a whole other topic, which is, machine learning for automating like crypto utility token market-making and arbitrage and so forth.

So that there's a lot of interesting ML applications in the crypto finance space, actually, you can do some things. There's already all sorts of AI in traditional finance. In the crypto domain, you can do some things you couldn't do in traditional finance because it could kill your characteristics there.

And that's fun. If you play with quantitative finance, it's like a whole new playground. But then the other thing I'm doing, which I wanted to save a little bit of time for is, which is what I'm spending probably half of my own time on now, which is the biggest chunk I'm spending on anything. So, we have one more spinoff entity from SingularityNET sort of an ensemble, which is called TrueAGI. And this is as the name would indicate my attempt to move back further toward my roots. And using the SingularityNET platform and the OpenCog system and so forth, make it a full-on push to actually get human level general intelligence .

This is based on what we're calling OpenCog Hyperon, which is a new version of the OpenCog AGI engine, and then building some specific learning systems while we build out OpenCog hyperon system. SingularityNET is then part of the plumbing, part of the platform for them. It lets you distribute pieces in a decentralized way, nicely.

**CRAIG:** And in that research, how do you divide your time?

**BEN:** Pretty haphazardly. In daily practice in the morning, I'm being a CEO and doing a seemingly endless series of meetings. In the afternoons, being a dad and romping with my two year old on the beach and stuff.

And then in the evenings, which can run quite late I’m doing AGI R&D stuff. Leads to dramatic sleep reduction scenario, but I get a lot of cool stuff done.

**CRAIG:** On the AGI research, It's not simply a matter of stitching things together. There has to be some basic breakthroughs in unsupervised or supervised or …

**BEN:** I think that the recent hype about unsupervised and self-supervised is pretty funny because some of us have been doing that for many decades.

And folks were doing that even before I was born, but now somehow people from a big company do this and it becomes exciting. And they're just like Yoshua Bengio, great guy, he published a paper on the consciousness prior oh, the human body and the three dimensional structure of space should be considered as part of distributions to inform probabilistic learning.

And I had a paper called the embodied communication prior like 2008 or nine or something, which said the exact same thing, but with a lot more detail, but nobody paid attention. But now a deep learning guru says the same thing very nicely, but with less detail into it and suddenly that's a revolution.

And then, it's the same with unsupervised learning. A lot of us never thought supervised learning had much to do an AGI. And now suddenly a whole group of people lead you down a completely useless direction for AGI, although, useful for many applications, then they're like, oh, actually, maybe we should do this other thing.

And then they get credit for proposing what their opponents were proposing for 50 years. So of course, labeled training data was never, was never the way to go for AGI anymore than expert production rules like they're using in the seventies were the way to go for AGI, nor is unsupervised learning as the mainstream is doing really the way to AGI either.

So, I guess next year, someone will rediscover active learning, like they'll rediscover that, that integrating action and perception is actually useful, then they'll rediscover like social learning as psychologists have known for a while. Then they'll get finally to embodied in interactive learning, which is shared with AI and other intelligent agents in the shared environment and developmental learning.

And then they will have rediscovered what many of us were publishing in the 1990s and eighties? Yeah, of course, going from labeled data to unsupervised pattern mining is a step forward, but it's not a big enough step forward. And I think that the core problem is that everyone working for a big tech company is implicitly, if not consciously, they're driven to work on things that will best leverage the unique economic resources of the big tech company. So, if you work for a big tech company, you are pushed to do things that exploit huge amounts of data and use processing power above all else, because that's the unique, competitive advantage that the tech company has.

And this, I don't think this is explicit on the point of view of most AI researchers. It's rather that the IT tool set that you have at your disposal, the APIs that you have are ones that differentially utilize these server farms and these datasets. And it's so much easier to script something that wraps up your proprietary APIs than to bother to do something else. I've tried a number of times to collaborate with some of my many close friends inside Google. And there's some bureaucratic obstacles to collaborating with outside open source AI projects. Those can be overcome. But the main thing is if you're in Google, it's just so easy to do things involving Google's internal proprietary API, and then so annoying to have to cooperate with people who can't use that API and who were writing code that doesn't use that API. so that totally makes sense. But what it means is that, when they talk about unsupervised learning, what they mean is mining patterns from the whole web or all of YouTube or Facebook videos like that.

They're looking at massive scale pattern mining to leverage the data that they uniquely have. And that, that can be very interesting for many practical applications, right? So, I don't discount it as a commercial thing at all. Like video analytics is barely touched, right? And because it takes so much compute time, but as compute gets cheaper and algorithms get better, there's going to be so much positive stuff and scary stuff that's going to be done with unsupervised machine learning on video and audio. But whether that is the path to AGI, I think that's the next multi-year AGI dead end the big companies will undertake. But the thing is unsupervised and self-supervised learning, clearly are part of the key. It's just the way you do that.

If you're trying to most rapidly get to the best possible unsupervised learning results on all of YouTube archive or something is different than the way unsupervised and self-supervised learning are used in like a young child's mind or something. So that's

**CRAIG:** that's right, because that data is already digitized. So, it's convenient.

**BEN:** And you don't have, you don't need active learning, right? So, there's a saying from Heinz van Foerster, a systems theorist, if you want to see, learn how to act, right? Because when a human or an animal is perceiving, the premotor cortex is heavily involved in low level vision for a reason, right?

When a human is perceiving they are often manipulating the environment also. This is why a lot of the time a kid who's immobile because of some medical reason, even if they can see they wind up artistic in some way, the perception-action combination, the synergy isn't there. So, what we're looking at with unsupervised learning now is it's like a massive unsupervised perceptual cortex, and then you can try to integrate reinforcement learning with it in some way, but reinforcement learning is its own perversion, which bears loose resemblance to the sort of more diversely motivated learning that a young child or that any of us does. So about whether we need a big breakthrough or not. That, I think remains to be seen, like what we've done with the OpenCog, is we've been taking deep neural nets and an evolutionary learning algorithm and which is evolutionary program learning with some probabilistic learning aspect, and then a logic engine.

We're making them work together on, on this knowledge graph in a sort of Blackboard type architecture. And we're using that underlying this clinical trial stuff. We've used it in simple ways behind Sophia, we're using it in some experiments in the Minecraft world and we're just hitting scalability issues with the infrastructure, basically. So unfortunately, we need to take a year to two years to rebuild that whole infrastructure of an OpenCog in a massively more, more scalable way. And so, making a massive distributed version of our knowledge hypergraph, we're making a new, gradually typed functional programming language to comprise the semantics of the nodes and links in the atom space.

And then basically reimplement all of our neurological and evolutionary algorithms in this more scalable infrastructure. So, if we are correct, this can have the same impact that GPU's did for deep neural nets. Cause there, of course there are many new ideas in the neural net literature, but the crux of it is about the same as in the 1960s.

And the multi-layer perceptron was a deep neural net. It just didn't work well enough back then. So, in the same way as if we're right about the sort of OpenCog paradigm, with this hybrid approach to AGI, if we're correct there, then getting a scaled up infrastructure will that thing start succeeding more dramatically and progressing more rapidly just as the advance of the Nvidia matrix libraries for GPUs did for deep learning.

That's not to say there won't be a hell of a lot of things left to invent along the way, but you wouldn't say since GPU's became cheap and prevalent, there hasn't actually been a huge mathematical or conceptual breakthrough in the neural net space. There's been a lot of really good breakthroughs. And so, it might be that after we have a scaled up infrastructure for OpenCog, then we get to AGI with a whole bunch of very good breakthroughs, rather than a paradigmatic shift, at least that, that's our working hypothesis, but we're a few years from validating that, but still that's exciting that it's not a few decades away from validating it.

**CRAIG:** On your timeline on the AGI . You were talking about scaling that, that up. When do you expect that to happen? What has to happen in order to scale it up?

**BEN:** Of course, as with any of these startup projects, a lot depends on how fast we ramp up funding into the project, right? Because we're not, this is not Google DeepMind. This is a, TrueAGI is a startup spin off of a decentralized blockchain based nonprofit, which is itself a startup. But assuming the resourcing works out, we've got a year or possibly a little more of just rebuilding the guts of OpenCog to make OpenCog hyper. We'll be doing some other AI, R&D, AGI R&D in parallel, parallel with that.

And then we start with a more interesting phase where we're doing a combination of building and teaching. And so, we're looking at multiple agents cooperating to build stuff and in a Minecraft like world, and then we're looking at Sofia type robots controlled with the TrueAGI system, but based more on, on embodied interactive learning where they're grounding, every, everything that they talk about in things that they can directly experience. And then I would say if the paradigm is right we're three to five years from a huge breakthrough in AGI. Of course, if it happened in, in two years, I'd be even happier, if it happened in seven years, I'm not going to shoot myself in frustration. But my friend Ray Kurzweil has projected 2029 for the breakthrough to human level AI just based on generic curve fitting type arguments. And my, my timeline is not extraordinarily far off from that. If it's five years we're beating Ray by a few years.

**CRAIG:** But on funding, are you talking about VC funding? Are you talking about national science foundation funding? How do you come up with the funding?

**BEN:** There's a wide assemblage of funding mechanisms out there? So TrueAGI is a for-profit company which can take traditional investment money. SingularityNET was it did an ICO in 2017 and the SingularityDAO, which is our decentralized finance spin off is creating whole new mechanisms of directing funding and interesting technology projects.

Yeah, we're, we're open to any and all methods of getting resources into projects that don't have the impact of handing over control to entities that we don't want to have control, right? Like I'm not incredibly interested in like creating a startup and then immediately selling it to a big tech company.

I think the whole decentralized AI methodology underlying SingularityNET, part of the idea there was, we're going to be better off if AGI comes about as A) open source, but B) with actually decentralized contribution and control and by building TrueAGI on SingularityNET platform, you're at least increasing the odds of that happening. But that, that goes into AI ethics and a whole other dimension of the topic that we haven't gotten into, right? Cause there's ethical risks with the vast unwashed mass of humanity collaborating in a decentralized way. There’re also ethical risks with a self-appointed elite group keeping proprietary, secret control and balancing those risks at the moment is more of an art than a science.

I'm somewhat, instinctively, and on a human basis, like going for the open and decentralized approach but you can't be certain about something like that in a scientific way. It's an unprecedented situation.

**CRAIG:** That's it for this week's podcast. I want to thank Ben for his time. If you want to read a transcript of this episode, visit eye-on.ai.

I find that reading transcripts reveal much that the ear misses. You should also check out SingularityNet at SingularityNET.IO.

Ben thinks the singularity is near. I do not. But regardless, AI is about to change your world. So, pay attention.