then it's so good to have you here it's fantastic to be here thanks chris now you told time magazine i want to understand the big questions the really big ones that you normally go into philosophy or physics if you're interested in them i thought to building a i would be the fastest route to answer some of those questions why did you think that well when i was it i guess when i was a kid my my favorite subject was physics and i wasn't sure know the big questions fundamental nature of reality what his consciousness or the you know or the big ones and usually you going to physics if you're injured and that but i read a lot of the great physicists some of my all time scientific heroes like fine men and so on and i realize in the last sort of twenty thirty years we haven't made much progress in understanding some of these fundamental laws so i thought why not build the ultimate tool to help us which is artificial intelligence and at the same time we could also may be better understand ourselves in the brain better by doing that too so not only was an incredible to was also useful for some of the big questions itself superstring so obviously i can do so many things but i think for this conversation i'd love to focus in on this theme of what it might do to unlock the really big questions the giant scientific breakthroughs because it's been such a theme driving you and your company yeah so i mean one of the big things i i can do and i've always thought about his where we're getting you know even back twenty thirty years ago the beginning of the internet era and computer era the amount of data that was being produced and also scientific data just too much for the human mind to comprehend in some many cases and i think one of the uses of a i used to find patterns and insights in huge amounts of data her and then surface that to the human scientist to to make sense of and make new hypotheses and conjectures says seems to me very compatible with the scientific method right the gameplay has played a huge role in your own journey and figuring this the on his is this young lad on the left there who is that is so that was me i must be about run nine years olds i'm captaining the england under eleven team and were playing in a for nations tournament's why we would wouldn't read think we're playing france scotland and wales i think it was not is so weird yet as that up and meters of it in my dreams it's i mean this is that okay so you were he it wasn't just chess you you are you loved all kinds of gay and i loved all kinds of games yeah and when you

launch deep mind meeting likley you started having it tackle gameplay white wine or look i mean games actually got me into a i in the first place because while we're doing things like we used to go and training camps where the england team and so on and actually back then guess it was in the mid eighties we would use the very early chest computers if you remember them as train on the train against as well as plenty of see each other and they were big lumps a plastic you know physical balls that used to somebody remember you stash you press the that the squares down and they were l e d lights came on an ironman the actually not just thinking about the chest hours actually just fascinated by the for that that this lump or plastic someone had program it to be smart and actually play chess to really high standard and i was just amazed by that and that got me thinking about thinking and how does this how does the brain come up with these thought processes these ideas and i maybe how we could mimic that with computers so yeah it's been a whole theme for for my whole life really but you raised all this money to to launch deep mind and pretty soon you're using it to do for example this so i mean this is an odd use of that what was going on here where we started awfully games at the beginning of the minds that back into any tends to this is for about ten years ago is our first big breakthrough because the we started off with classic atari games from the nineteen seventies the simplest kind of computer games there are out there and one of the reasons we use games is they're very convenient to to to to test out your ideas and your algorithms that really fast a test and also as your system's get more powerful you can choose harder and harder games and this was actually the first time ever that our machines surprised us the first of many times which it figured out in this game called break out that you could send the ball round the back of the wall and actually have been much safer way to knock out all the tiles with the walls as a classic atari game that that was our first we'll ah ha moment so this was not programmed to have any strategy it was just told trying to figure out a way of winning you just moved back yet about it was a fifth you can find a way of winning right it was a real revolution at the times is and twenty twelve twenty thirteen where we cut we coined these terms deep reinforcement learning and the keith thing about the knows that those systems were learning directly from the pixels the war pixels

on the screen that they want been told anything else so that we told maximize the score here the pixels on the screen thirty thousand pixels you gotta the system has to make sense on it's own from first principles what's going on what is controlling how to get points and the other nice thing about using games to begin with they have clear objectives right to win to get scores so you can kind of measure very easily that your systems are improving but there was a direct line from that to this moment a few years later wow their country of south korea and many other parts of asia and in fact the world went crazy over of a lot yeah so this was the pinnacle of this is and twenty sixteen pinnacle of our games playing work were so we done atari week we'd we'd done some more complicated games and then we reached the pinnacle which was that game of go which is what they play in asia instead of chess but it's actually more complex than chess and the the actual brute force algorithms that will used to kind of crap chess were not possible with go because it's a much more pattern base game much more intuitive gay so even though the blue be carry caspar of in the nineties it take took another twenty years for our program alpha go to beat the world champion ago and we always thought myself in there and the people working on this project for many years it if you could build a system that could be the world champion it go would have had to have done something very interesting and in this case where we did with africa or is it it basically loan for itself by playing millions and millions of games against itself ideas about going to write strategies and in fact invented his own new strategies that the go wild had never seen before even though we've played go for more than you know two thousand years the oldest board game in existence so you know is pretty astounding i needed it when the match it also came up with brand new strategies and few continued this with with a new strategy of not even really teaching it and the thing about go but just setting up systems that just from first principles would would play so that they they could teach themselves from scratch go or chest talk about alpha zero and the amazing thing that happened in in chess yeah so following this alpha go started we start with africa by giving it all of the human games that been played on the internet so it started that as a basic starting point for it's knowledge and then we wanted to see what would happen if we started from scratch from literally random play

so this is what alpha zero was that's why it's the zero in that in the name is it started with zero prior knowledge and the reason we did that is because then we would build a system that was more general so alpha go could only play go the alpha zero could play any to play again and it did it by playing initially randomly and then slowly incrementally improving will not very slowly actually within the course of twenty four hours going from random to better than world champion level and so this is so amazing to me so i'm i'm more familiar with chess than his go and thought decades thousands and thousands of ai experts worked on building incredible chest computers eventually they got better than humans you had a moment a few years ago were in nine hours of the zero taught itself to play chess better than any of those systems ever did yeah to talk about that they are it was pretty incredible moment actually so we said it going on chess and as he said this is rich history of chess and a i were diabetes expert systems that been programmed with these chests ideas chest algorithms and you start you have this amazing i remember this directly we sort of sit down with the system starting off random you know in the morning you go for a cup of coffee come back i can still just a our beat it by lunchtime maybe just about and then you let it go if another four hours and buy dinner is the is the greatest chess playing entity there's ever existed and you know it's quite amazing like looking in that live in on something that you know well you know like chess and your expert in and actually just seen that front of your eyes and then you extrapolate to what it could then do in science or something else what if which of course games were any a means to an end they were never the end in themselves they would just the training ground for our ideas and said make quick progress in a matter of you know less than five years actually went from atari two to go you know many this is what people are in or of a i also kind of terrified by and it's not just incremental improvement the fact that in a few hours you can achieve what millions if humans over centuries have not been able to achieve that that it's just that gives you pause for thought i mean this is us i mean it's a hugely powerful technologies can be incredibly transformative and we have to be very thoughtful about how we use that capability so talk about this use of that because this is again this is another extension of the the what you've done one now you're telling it something incredibly useful

for the world and what what are all the letters on the left and what's in the right yeah so this is our this was always my aim with i from from a kid which is to to use it to accelerate scientific discovery and actually ever since some doing my undergrad at cambridge i had this problem in mind one day fat for a ice called the protein folding problem and it's kind of like fifty a grand challenge in biology and and this race inputs explain proteins are central to life that the building blocks of life everything in your body depends on proteins and you describe your protein protein sort of described by it's amino acid sequence which you can see and covers roughly the genetic sequence describing the party each other that's what the letters and each of those letters represents in itself of a complex molecule that's right each of those letters and amino acids and you can think of them as a car you can think of as a kind of string beads there at the bottom left right and and but in nature in your body or in an animal this string a sequence turns into this beautiful shape on the right that's the protein that's that those letters describe that shape and that's what that's what it looks like in nature and then courts and thing about that three structure is the the three the structure the protein goes a long way to telling you what it's function is in the body what it does and so the protein folding problem is can you direct you predict the three d structure just from the amino acid sequence so literally the give the machine the i system the letters on the left can produce the three d structure on the right and then that's not alpha dogs outputs are those it's not calculating it from the letters it's looking at patterns of of the folder program proteins a known about and i saying and somehow learning from those patterns yes maybe the way to do so so when we started this project as you stray off to africa oh i thought we were ready once we once we wieder crack go i felt we will finally ready after you know most twenty years of working on this stuff to actually tackles and scientific problems feeding protein folding and what we start with his painstaking over the last forty plus years experimental biologists have pieced together around one hundred and fifty thousand and protein structures raising very complicated you know x ray crystallography techniques another complicated experimental techniques and the rule of thumb is that it takes one phd student that whole phd so for five years to uncover one structure but there are two hundred million proteins knowing to nature so

you could just you know which take forever to do that and so we managed to actually fold using our four fold in one year all those two hundred million protein known to science so that's a billion years a phd time it's it's amazing to me just how reliably it works and on the and that the shows you know here's the model and you do the experiment and sure enough the protein turns out the same way times and two hundred million yeah and the what and the more you deeply you go into proteins easy you just start creating high school visit they are me look at how beautiful these proteins on each of these things do a special function in in nature and they're almost like works of art and still astounds me today that awful can predict the green is the ground truth and the blues the prediction how how well it can predict is to within the width of an atom on average is is how accurate prediction is which is what is needed for biologists to use it and for drug design and for disease understanding which is what will our for photon locks humid a surprising decision which was to give away yet the actual results of your two hundred million per another we open sourced for phone and gave everything away on a huge database with our wonderful colleagues european by informatics institute i mean your your part of google with that was there a phone call say this what what did you just do well we we know i'm lucky we have very supportive google really supportive of science and and understand the benefits this can bring to the world's and you know the argument here was that we could only ever have even scratch the surface of the potential what we could do with this this you know maybe like a millionth of of what the scientific communities doing with it there's there's over a million and a half biologists around the world have used our for phone foreigners predictions we think that's almost every biologist in the world is is making use of this now every pharma company so we'll we'll never know no probably what the full impact of it all is but you are continuing this work in a new company that spinning out and s google code as some of islamophobia a similar thing give us just a sense of the vision that was division so awful fold a is is a sort of fundamental biology to like how do you know what are these sweetie structures and then what might they do in nature and then if you you know the reason i thought about this and was so excited about this is as is that this is the beginnings of understanding disease and also may be helpful for for designing drugs so if you know

the shape of the protein and then you can kind of figure out which part of the surface of the protein you're going to target with your drug compounds and ice said morphic is extending this what we did it now falls into the chemistry space where we can define design chemical compounds that will bind exactly to the right spot on the protein and also importantly to nothing else in the body so it doesn't have any side effects and it's not toxic and so on and we're building many other ai models sort of system models to alpha fold to help predict make predictions of in chemistry space so we can expect to see some pretty dramatic health medicine breakthroughs in the coming to the i think we will be able to get down job discovery from years to maybe months okay dennis i'd like to change direction of it a mutual friend the worry gave a talk last year it today i that she called the moloch trap the might trump is is a situation where organizations companies in a competitive situation can be driven to do things that no individual running those companies would buy themselves do and it's it's felt i was really struck by the sort and it's felt as a sort of lay person observer that the molotov top has been shockingly in effect and in the last couple years so here you are with deep mindset of pursuing these amazing in medical breakthroughs in scientific breakthroughs and then suddenly kind of out of left field open eye open a i with marks of release his church cpt and well goes crazy and send because holy crap ai is is you know everyone can use it and and there's a sort of them it felt like the up trap and actually i think marks of ceo such in that are actually said google is the a hundred pound gorilla in the search space we wanted to make google dogs how and it did google did not there was a dramatic response your role was changed you took over the whole google ai effort products so rushed out gemini some part amazing part embarrassing i'm not into law school gemini because you've addressed elsewhere but but it's it feels like this was the moloch trap happening that you and others were pushed to do stuff that you wouldn't have done without this that of catalyzing competitive thing metre did something someone as well they rushed out an open source versions of ai which is arguably a reckless that in itself the seems terrifying to me why is it terrifying look it's a complicated topic of course and first of all i mean that many things to say about it first of all we be we were working on many large language models and fat or with the google

research actually invented transformers as you know which is the architecture that allowed this to be possible five six years ago and so we had many large models internally the thing was i think what the church u p t moment did that change was and fair play to them to do that was they demonstrated i think it's somewhat surprisingly to themselves as well that's the public we're ready to you know the general public are ready to embrace the systems and actually find value in he systems impressive though they are we i guess when we're working on these systems mostly you're focusing on the floors and the things they don't do and hallucinations and things you're familiar with now we were thinking you know would anyone really find that useful given that it's does this and that the other and we wanted to improve those things first before putting them out the interesting the it turned out that even with those flaws many tens of millions of people still find the very useful and so that was an interesting update on maybe the convergence of products and the science the actually all of these amazing things we've been doing in the lab so to speak are actually ready for prime time for general use beyond the rarefied world of science and and i mean that's pretty exciting in many ways so the million and we've got this exciting arrive of speak which were all enjoying enjoying a or this jenner today i stuff is amazing but let's roll the clock forward of it microsoft and open a i reported to be building are investing like a hundred billion dollars into an absolute monster database super computer that can off the computer orders of magnitude more than anything we have today i think the is it takes like five gigawatts of energy to drive this it's estimated that the energy of new york city and to drive the data center so we're pumping all this energy into this giant va spray google i presume is going to match this type of investment right what we yeah i mean we don't talk about a specific numbers but where you know i think we're investing more than that over over time so reasons we we teamed up with cable back and twenty fourteen is we we were kind of we knew the in order to get to a g i we would need a lot of computer and that's what's transpired and google you know had and still has the most computers so since it's earth is building these giant computers that are going to base is giant brands that are going to power so much of the future economy and it's all it's it's by companies that are in competition with each other how how we avoid the situation where

someone is getting a lead someone else has got one hundred billion dollars invested in their thing isn't someone going to go wait a sec if we used reinforcement learning in here to maybe have the ai tweak it's own code and we right itself and make it so how we might be able to catch up in nine hours of the weekend with what they're doing roll the dice i'm at we have no choice otherwise we're going to lose the forgotten for our show us how how are we going to avoid that yeah well we must avoid that of course kaylee and my view is that as we get closer to a g i we need to collaborate more and neck the good news is that the most of the scientists involved in these labs know each other very well and we talk to each other a lot at conferences and other things and this technology still relatively nation so probably it's okay what's happening at the moment but as we get closer to a g i i think as as a society we need to start thinking about the types of architectures that get built so i'm very optimistic of course that's why i spent my whole life on on am working on a i am working towards a g i but i'm i suspect there are many ways to build the architecture safely robustly reliable reliably and in an understandable way and i think there are almost certainly going to be ways of building architectures that are unsafe or risky in some form so i see a sort of what kind of bottleneck that we have to get humanity through which is building safe architectures as the first types of a july systems and then after that we can have a sort of a flourishing of many different types of systems that perhaps shard of those safe architectures the are ideally have some mathematical guarantees or at least practice guarantees around what they do to do governments have an essential role here to define what a level playing field looks like and what is absolutely taboo yeah i think it's not just about actually added government and civil society and academia and an awful positive side to have a critical role to play here to shape along with industry labs what that should look like as we get closer to a and they corporation needed and a collaboration needed to prevent that kind of runaway race dynamic happening okay with the sense that you remain optimistic what's what's this image here it's one of my favorite image is actually i call it like the tree of knowledge so you know been talking about science and a lot of science can be boiled down to you know if you imagine all the knowledge that exists in the world as a tree of knowledge and then maybe what we know today as

as a civilization is some you know small subset of that and i see i as this tool that allows us as scientists to explore potentially the entire tree one day and we have the this idea of route no problems that like how for followed the protein folding from way wife you could crack them it unlocks an entire new branch of discovery or new research and that's what we try and focus on a deep mind and you would eat mind to crack those and if we get this right then i think we could be you know in this incredible new era of radical abundance curing all diseases spreading consciousness to the stars what you know the single and human flourishing work out of time but what's what's the last example of like in your dreams that this dream question that you think there is a shot the of your lifetime a i'm i might take well i mean once a g i bill what i'd like to use it for ways to try and use it to understand the fundamental to nature of reality say duke's parents at the plank scale you know the that the smallest possible scale theoretical scale which is almost like the resolution of reality you know i was brought up religious and in the in the bible there's a story about the tree of knowledge that doesn't work out very well is this is is the is there any scenario where where we discover knowledge that the universe as humans you may not know that potentially i mean there might be some on noble things so but i think scientific method is the greatest sort of invention humans have ever come up with you know the enlightenment and scientific discovery das was built this incredible modern civilization around us and all the tools that we use so i think it's the best technique we have for understanding the enormity of the universe around us well done as you've already changed the world i think probably everyone here will be cheering you on in your efforts to ensure that we continue to accelerate in the right direction and kick them as has out of thank you for your