UCL Brain Stories Episode 18 - Brain Stories Live

Wed, Nov 08, 2023 10:21AM 1:35:47

SUMMARY KEYWORDS

ai, work, neuroscience, brain, study, models, problem, people, fact, machine learning, data, learn, understand, ucl, neuro, question, caswell, metacognition, architecture, decision

---

00:02

Welcome to brain stories live. I'm Castleberry This is Lena ray.

00:07

So this is new. This is new to us is to you who's seen or heard our podcast?

00:15

reasonable numbers. Okay. So in the podcast and tonight, we've got pretty much the same format. First of all, we talk about the science or people are doing what I guess are doing, what they're researching. And then we talk about what makes them tick the journey that brought them to where they are, and we're going to be doing the same thing tonight.

00:31

By way of audience research, who we got here we've got anyone any undergraduates here, show of hands, no, undergrad, I saw you there. No.

00:41

Any any postgraduates? few of those? Anyone who's not at the University at all, feel it. Okay. Good. Good mix, right. Fantastic. Now, now we know who we're aiming for. So Celina for a moment. Well, welcome, everybody. And thank you for coming to join us for this experiment. And so if Caswell hinted at this is there's a series of firsts this evening. It's our first ever event of recording where we've had more than one guests. So we will welcome a panel up to the stage in a moment. And it's also our first ever live recording, hopefully not our last, please be nice to us, so that we come back and do this again. And the reason we decided to do a live recording is because we wanted to meet some of our listeners, but also because this year is the 50th anniversary of the UCL neuroscience domain. And so the UCL neuroscience domain is a network across UCL that tries to connect scientists working in different disciplines bring us together, so that we can work more effectively, but also to do outreach events so that we can showcase our work more broadly. And so you're all very welcome tonight, and we're really looking forward to sharing some of our amazing scientists with you. And we will, the way the evening will work is we will do this in two parts. The in a second, we'll welcome our panel up. And in the first part, we will discuss their research where they think their field is going and what they're excited about, will then take a short interval in about 15 minutes time. And after the break, we'll get a little bit more into the start the scientists story. So why did they become interested in their research areas? What did they study? What key career decisions did they make that brought them here? We really want this to be interactive. So you're very welcome to ask questions throughout, just give one of us a wave and we will come to you. And if people have questions, but they don't really want to stick their hand up, I've got as you can see, in my hand, I'm clutching a few pens, and I've got a little bit of paper and I'll leave them down at the front. So you can write things down in the interval, and then we'll ask them at the break. So I'll hand back to Caswell, we have an organising theme today. So the topic, the thing that unites three guests, is neuro AI. And so what I'm going to do is attempt to define what that is now, I'm sure when I guess Come on, they'll tell me I'm wrong. And that it's nothing like that. So here's my working definition of neuro AI. So it's a portmanteau of neuroscience and AI. The point being that these two fields sort of share a common lineage. So there's sort of a natural connection between them. And what that means is that various points, there's been sort of exchange of information. So some of the ideas that are sort of driving AI development now been borrowed from neuroscience, the ideas of neural networks, the ideas of reinforcement, learning the way animals learn, things like convolutional networks, which is how we think some of the visual system works. Has that's been happening for a while. But what's increasingly been happening over the last sort of five or 10 years is, neuroscientists have got kind of wise to this and have been taking lessons from AI. And so neuroscientists are increasingly using machine learning tools to deal with their data, maybe to diagnose diseases, but also using these sorts of neural machine learning models as models of the brain telling us, you know, if there's a problem, how ought this be solved? And then we go and look in the brain to see if the brains respond like the machine learning models. So we're going to find out whether the guests agree, you've heard enough from me already. So I'm going to say the magic word. Let's bring on the guests. And hopefully three eager researchers will bust through the curtain.

04:28

Fantastic.

04:33

Well, thank you for joining us this evening. And we're very excited to have the discussion with you tonight. Perhaps we can start just quickly by each of you introducing yourself telling us your job title and where you are based in the infrastructure of the university. And Rick, I'll start with you. Okay. Hi, I'm Rick Adams. I work in the centre of medical image computing, computer science, but also in neuroscience in the Institute of coding.

05:00

neuroscience. And also one day I'm a consultant psychiatrist. I do clinical work and see patients as well. Busy man. Yeah, teacher. I'm Deeksha Gupta. I'm a senior research fellow at the

Sainsbury Welcome Centre. I sort of said between systems and computational neuroscience, and yeah.

05:19

Benedetta Hello. Hi, I'm Benedetto de Martino. And I'm a cognitive neuroscientist working at the ICN. That is the author of cognitive neuroscience, not much fun to see you in.

05:32

So maybe we can start just by giving the floor to each of you for a few minutes to expand, maybe introduce your research in in kind of general terms. What are you working on? What sort of techniques are you using? And why is this important? And so if we go in reverse order this time, Benedetto would you like to start? Sure, I do. Mostly my work is in human cognitive neuroscience. And my hi all is strange. I mean, I come from molecular biology, but I haven't seen like a little mouse for a long time besides, this is for the occasional one.

06:11

So we do use neuroimaging methods like fMRI, E, G, M, eg, but mostly my lab is also interested in computational modelling, and things that are like crossroad between neuroscience, economics and machine learning.

06:31

I've just completed a PhD and during my PhD I was studying how different brain areas especially like corticostriatal circuitry, coordinate the activity to do this, like very fundamental computation of how do you evaluate different pieces of like sensory evidence? And how does that decision process evolve? And finally, how do you come up with a decision that you want to take in this real world? And now, as a postdoc, I am developing a paradigm this study, compositional generalisation.

07:02

compositional generalisation is basically how do you use bits of your past knowledge to compose sort of intelligent seeming behaviour. So for instance, this is like a whole spectrum of behaviours, which falls under this, so on like, the simplest thing, and would be, if you've picked up some chords, then you can pick up a new song very easily, because you can use those bits of knowledge you've learned in the past, compose them together, and pick up a new song very easily, and under more complex cognitive, and would be that if you learn a few words, then you can use them in new contexts and new sentences to convey all sorts of meaning. So that's what I'm sort of trying to say the hippocampus different interactions during this sort of behaviour. Now, I enjoyed the use of the phrase seemingly, and

07:48

that's might be one for us to pick.

07:52

So, I work in this kind of subfield called Computational psychiatry. And there's a mixture of things that goes on there, I've done bits of all of them. So sometimes using computational models of cognition, to work out how my brain processes proceed, just like these guys do, but then also to think how they might go wrong. So how perception might become a hallucination, for example. And then also biophysical models of imaging data to try and infer neurobiology and neurobiological properties of the brain,

08:34

which is, so to link to things like drug targets, and that kind of stuff. And then lastly, kind of more machine learning based methods to analyse very large datasets, for example, comparing patients and controls and this kind of thing. So I just wanted to pick up on some of the things you were saying, Rick, so would you say it's true that we're already at a point where machine learning models are useful in the clinic? Or is that is that something that's yet to come? I'd say it's very much depends on which clinic so so in some areas, they're really kind of ready to go, it looks like from how well they work. So ophthalmology, there was a really landmark collaboration between DeepMind and the Institute of Ophthalmology down the road, where they had a million labelled retinal image scans. And this network learned how to diagnose diabetic eye disease, hypertensive disease, and tell the age and sex of the person in a way that they don't know how it can do.

09:38

But when the problem is

09:40

brain data, like structural MRI data or functional MRI data is such.

09:47

So many more data points in there. There's like 100,000 data points instead of the number of pixels in a retinal image.

09:56

And we don't have a million labelled scans we have like maybe

10:00

1000 or 2000? It's really nowhere near as as good as that in psychology. So some clinics, yes, dermatology, lots of Radiology, ophthalmology, watch out, but not psychiatry just yet. I know you optimistic or do you think this is the sort of lack of labels and lack of clear distinctions just doesn't mean you, we don't, we don't know enough yet to train a model to predict these things or

10:25

so I think, for the short, medium term, the way you would get results out of those kinds of methods is using a more hypothesis led approach, and maybe trying to, you know, discriminate between

10:43

groups who, who, you know, respond to one treatment or another treatment. And if you if you haven't, if you can collect enough data, and you know, in animals, where those receptors are, you can tell the machine to only look in the brain where those receptors are, for example, and reduce the dimensionality that way, something like that, you know, but to just chuck everything in and hope for the best, I don't think it's going to work until you have millions of scans. I'm keeping an eye on my watch, because I know, we'll probably chat all night otherwise, I I think this is really fascinating area. But I wondered if just for the benefit of anyone in the audience, who is maybe completely new to the concept of neuro AI, if one of the panellists could kind of just break it down a little bit what we mean when we talk about things like learning datasets, machine learning, and then how that can be used to then look at new datasets. So just really the kind of fundamentals of some of the things that we've just been talking about.

11:44

Okay, I'll give it a go. And,

11:48

you know, I've been working a long time in a field that was called neuro economics. And that seems that seems almost the trick is put the new row in front of another word, and then try to figure out what that means.

12:02

So the thinks is the way in which I said and there's some things we'll get used to now because there's been this kind of explosion of AI is just methods that we add in the field that they just got much, much more sophisticated. They're like, you know, statistical learning methods and so on. But the new things is the fact that that unlike

12:28

other type of engineering stick approach in which you as an engineer, you

12:33

come into engineering what the things has to do er, your engineering that the learning architecture, so your practical your engineering, not towards, in the case that

12:46

it was talking about, they didn't engineering something specifically for the AI is engineering, some architecture that it doesn't care if it's an AI, images, whatever, like a scan from an airport. And then through these like massive training, normal is a training through labelling. So this was the reason why it was saying, the problem is we do not have enough labels, can the machine start to understand this association, this relation? Now this is something quite a bit different from the way you learn and small children learn. So in a way, the big challenge Hi there is getting this massive data or understanding something's more about human learning habits, small children, I mean, they're not so small anymore. For adult they're also small, but when they learned were horses, they didn't need the million horses a billion horses to detect well horses, you show them for horses, and since then, they can pretty much detect where horses so is a learning architecture. Now neuro AI specifically

14:02

is clearly as we joke not super clear what it is that everybody sees in one way, but you can imagine both way like is the use of AI. In neuroscience. You know, Caswell, for example, has done himself some work he should be talking about that rather in interview, in which he has like, he trained and that machine to just detect the like movement of mice and things and, and doing like some work that was very tedious and very long for human to do and not very well. And then the most ambitious part of neuro high that unfortunately, is one as under delivered, in my opinion, is the contribution of neuroscience to do exactly what I told you to have a computer that is like my daughter, they see four horses and learn about horses. That was a big ambition of mine, and actually follow it very closely because I was doing my PhD together with demister

15:00

service at UCL. And you know, the big dream at the beginning of things like the mind was, let's have neuroscience respire us now. 10 years. Hide the line? Not sure as much, you know, you can see SPIRATION in everything. I mean, the idea of neural network is biologically sparse. But maybe at the moment has gone more in the other way has been more for neuroscience using this tool was satisfactory or not super. And I wonder if I might come back to you rageous. With a question about using AI in diagnostics, I guess you talked about the need for really big datasets and the availability of a sufficient number of scans to be able to kind of identify these patterns, what are the risks that because you I could imagine a scenario where,

15:54

you know, a machine learning or AI looks at a scan and says, Oh, this is abnormal? And it means this, but actually, could it be that the scan has been taken in a different centre on a different machine? And can we are we at a point yet where we can really distinguish those things? Yeah, so the, I guess, these issues, make it into the general media, they've obviously super important and of great concern to lots of people. And the basic problem is the the machine is only going to learn from what it sees, and literally just from what it sees. So the problem is

16:38

the extent to which it can generalise that we take generalising knowledge for granted, because we're so good at it. But computers are not necessarily so just like you say, it can learn from extraneous details that he's not supposed to learn from. So for example, there was one case of a

16:57

programme learning to diagnose pneumonias from chest X rays. And they discovered that it was it was working because it learned to diagnose the pneumonias, because the people who were more severely ill had ECG leads on and

17:15

because they didn't have time to take them off. And that was, and that was how diagnose pneumonia. And there are lots of other examples like

17:23

I think I've read somewhere but I can't so don't quote me on this as I can't remember the details that

17:30

one of these cool either ophthalmology or dumb. I think it's an ophthalmology, G deep net that can diagnose these problems that was talking about was then subsequently tested on a whole bunch of retinal images in India, I think possibly and and it didn't work. And so, you know, there have there has to be

17:52

training set data from all over the world from all kinds of different scanners if it's going to be exported, because obviously, the hope is we can use this technology in places that a resource poor, but obviously, the absolute worst thing is gonna happen if it's if it doesn't work in those in those places. So.

18:10

So there's a, there's a really good point, actually, I want to come back to those sort of questions of generalizability. And sort of how equitable we can use these things. But just want to finish with one finish up on one of the points that Benedetto mentioned, which was this sort of disappointment that the dream of neuro AI to sort of paraphrase was that the information would go both ways. And at moment, neuroscience is doing pretty well. But some people would claim and indeed, when I put on your AI meetings, the lack of machine learning, people would seem to indicate that actually, the information isn't going back the other way, so much. And then, an example of sort of a story that was told me was, you know, it's a bit like birds and planes and flying. If you want to, if you want to build a flying machine, it's great to see birds at the beginning, because they tell you that flying exists. But if you didn't spend 50 years trying to build something with floppy feathery wings, you're not going to get very far you need to give up and try something different. And the implication being that maybe we've taken the inspiration from the brain, but actually, we shouldn't be trying to copy it too closely to achieve what we want. Do I just wondered whether you've got any sort of thoughts or thoughts about that, whether whether you're disappointed or not, basically, I'm more optimistic. I think from a basic science perspective, I think the fact of the kind of techniques which have been developed in AI of late have helped us make better models of like, how neural responses look in the brain. But I think now is the time that we can take that inspiration from neuroscience back to AI models. So AI models tend to be really brittle as you were saying that they don't generalise well, they pick up on random features, whereas brains tend to learn functions which are very smooth and generalise well and are not brittle in the same way. And I'm optimistic that some of the properties biological properties of the brain and that there are excited very excited Tory and inhibitory neurons. There were certain projection pattern. The architecture looks a certain way, all of them

20:00

these properties, if you like, explored them, we would come up with something which is more robust in some way. And there's like inspiration to be drawn from those kinds of biological features, which might help build more robust systems, which are also more energy efficient. Hopefully.

20:16

That's what I think I can actually

20:20

reply to you as well. I would have shared these things until probably six months ago. And and the things is, after the explosion of general language model, what has happened, a big shift in the field that I think, being in academia, we don't know this so much. And the fact that at the beginning, the aspiration was this really long term goal. And as I agree with you, neuroscience can be really useful to achieve this long term goal. But now the economic pressure after you know, opening, I released church GTP has been so strong, that now a lot of effort in this company, is now going to this general language language model that are very far away from the brain. And in a way there is also good news in that for academia means that there is a lot of work for us left.

21:16

But it's also a bit sad that the idea that

21:21

at the beginning, there was this idea now, the best of science and things was the dream of a lot of people, including people, probably me being pessimistic now, maybe things will change. But the problem is, and I think is almost this as a political problem, if we delegated the study of this thing to industry, or being the industry will do, what we should do as academia without spending money, and they'll do for us, there is a big risk because the industry overreact, react very quickly. And now industry is overreacting, in my opinion to this general language model, was the same reason when there was a point in which Google was scanning all the book for the making these digital access, and a lot of university withdraw to finding funding these things. But the problem is that Google hasn't need to do it. And at one point, they decide, we stopped to do it. And my wife that is an archaeologist, I remember at the time, she was oh, crap. And now Nope, there are no funding for doing that. So I think this is actually a really good point in which, because it could be a really nice having a conversation with industry, but I think academia has been sitting on the bottom too much in hoping that, like industry will do instead of us. But as we can see now industry is might be taking a completely different path. And that's Caswell say, might be interested in knowing bird anymore, might be interested in plane, but we shouldn't really hope and now we'll understand bird biology, but Boeing and we complain with Boeing that he's not helping us with bird biology. So that was my only. I mean, I think I agree. Like, I think engineering goals are different, and they're gonna go a certain way. But if the real capitalist gains to be had by using this more efficient architecture, if you can find it through the brains, then they probably will co opted. But yeah, probably diverging in that way. Yeah. Because I mean, maybe it's worth one of you saying why why you think things like the sort of recent success? Or what transformers are, what they've, what successes they've driven? And maybe most importantly, why? Why neuroscience tend to think they're not good models, the brain Oh, some not everyone agrees. I'm on the fence. But Would one of you like to say something about that?

23:37

I mean,

23:40

I can spit Yeah, I actually yeah. So transformers I find quite interesting, because I don't know much, I don't know huge amount about them. But they essentially

23:51

they're like a kind of short term memory unit in a in a in a neural network that allows the network to kind of store information and let it reverberate for a short time.

24:04

While an input a stream of input is coming in, and and they have really revolutionised

24:11

the language models and speech recognition models, because those models, unlike spatial

24:20

can visual models that need to integrate information over space, to understand speech, and to be able to predict the next word, which is how they train these models. And you need to integrate information over time. You can't just use the last word you need to leave us the last few sentences. And, and the question is, does is this how the brain predicts speech? I mean, probably not exactly like this. But it definitely must do it something like this. And there's there's a really interesting angle I think,

24:50

related to computational psychiatry, but related to this, which is

24:55

the in schizophrenia the commonest kind of hallucination is auditory verbal

25:00

hallucination I get that you can have all kinds of hallucinations, all kinds of modalities, but voices are the most common. And the weird, it's a very strange thing that why is it specifically voices when you could hallucinate anything. And one reason may be that if if there is, for example, dysfunction in receptors that help us integrate information over time, like NMDA receptors may be that function that is, is really impaired in some people, and that particular function starts to degrade before others do and then, and then you become much more reliant on your expectations of that model, rather than the information that's coming in, and then you that can generate hallucinations. But you only really think about these properties when you have to build these networks that can recognise speech or recognise facial things. And you realise what is so important about the differences of it also are like one of the major departures in transformer architectures is that earlier, when people were trying to predict speech or like sequences and time, they were using recurrent neural networks, which were probably like, closest thing you can get to what brains look like in like artificial neural networks. And the advance in Transformers is that they got rid of that component and said, We do not need that we can just do with these matrix multiplications. And these attention heads. So it's like a major departure from what looks like brain like, artificial units to like a completely different thing. All the people are trying to map that onto an RNN recurrent neural network, but it's like that mapping is still incomplete and debated.

26:38

Great, thank you. So I want to move on. Now to go back to the points you're mentioning earlier and pick up on those about the sort of

26:48

about the risks inherent with these approaches, whether applied in a clinical or even in academia, whether we're sort of being led down a blind alley, or be misled about the way the brain works.

26:59

Rick, you've already sort of said sort of the the necessity of making our training data look like the things we actually want to test. So I think your example was, you know, if you could just train on eyes in Western Europe and hope that your network is going to work somewhere else, then actually, you're basically trusting blind faith, there's no reason to go. I went with one of the other two of you would like to sort of whether you see there any risks for us, both the scientists or as, as practitioners and going this way, and maybe I'll give you a sort of, I don't know, a line to sort of start with. So I was talking to someone recently, who said, Oh, I'd never want an AI doctor, because I don't really know why they made their decisions. And my immediate response was, Do you know what your doctor major decisions? They could, you know, they could just say anything? And you'd be like, oh, yeah, it's the it's the weird wibbly thing. Thanks, doctor.

27:47

But I think here the point your friend, I think, and it's actually quite an important point, I'd be now recently talk with a colleague at UCL law. And the fact is, we live in a society in which we need to be able to at least justify our decision. So if you do an accident, you do something and you turn right and you make a decision and you kill somebody, when you go in court that people will ask you love, a lawyer will ask, Why did you do that, and you're trying to give them motivation. We have a society in which intention intentionality matter. Because if I say I did, because I wanted to kill that guy is very different from idea to just. So this is an issue that is actually incredibly important. In human brain architecture, we have this module that I happen to study quite a bit, that is called meta cognition, and is the fact that we are able to introspect what we do and verbally report. This module hasn't been very much of interest for AI, because even for human you wonder, why do I even need that module. And, you know, we can argue that we're social animal won't communicate, but see about things like what your friend was telling you. The fact that that if a machine the something's wrong, and doesn't detect the bomb, or whatever, because he has no access into the decision process, the tingle has happened in the neural network, in a such high level that in the findings, there will be no engineering given because the best theory is that these things work, but we don't know where they work and how they work. And the fact that you can't just You're right, maybe your doctor will come up with something. But we live in a society in which we not only do things, but we also introspect what we have done and communicate with each other. And I think this is going to be a massive problem. If we carry on these EIR, modern architecture, that will be for legal reason, even if those cars are going to be better. It will be very hard to use them because when then they need to go to court, who you're going to be

30:00

into court. And in the machine won't even be we have just ignoring all these parts of the architecture that is able to introspect. So I think is more like a, you know, it seems funny, but actually, I think I would agree with your friend that wasn't

30:21

Caswell has got no

30:24

follow up on one of the Sony related study I saw today. And I've seen there's several out there when, when people have tried to implement

30:35

deep learning and other AI technologies in the clinic, which seem to work super well like the ophthalmology, one, radiology ones.

30:46

When they work better in many cases than consultants doing that job, and they can do it 24 hours a day.

30:54

30:54

They've often found that in the lab, they outperform the humans, but in the clinic, they're no different. And when they look into it, the reason one of the reasons they're no different is because when they give a recommendation the doctor would not have made, the doctor just overrules it, and ignores it. And so really, the the this reasoning, this big ability to present its reasoning is probably going to be the only way that actually gets people to do the recommendation to follow it. Because it may not actually make sense rationally. But practically, they might not be useful unless they can do that. So my sort of take home, I guess from that little bit of the discussion is we are not in danger yet of having our GPS replaced by a or do we say chat GP?

31:50

I get that wrong.

31:54

Yet.

31:56

But maybe that can lead us into a bit of a discussion of, I guess, overall, what we're seeing is there's areas of huge potential, but it's still a really early field where there are huge challenges to overcome. So I guess from each of you, I'd really like to hear a minute or two about what you think the next 15 years holds for Neuro AI in your particular areas. What are the things that you're most optimistic and excited will happen?

32:26

Starting with me? Yeah. Okay.

32:28

So I think I'm most excited about neuroscience has worked with very like, sensory and motor systems. And that's where a lot of the research has been sort of confined, and we use simple models explain that. And that was all well and good. But now we're moving into this like, era in which we're trying to understand more complex decision making and behaviour in general. And I think currently, AI is the only field which has models for that. And I'm very excited about marrying those ideas with like, what we know happens in the brain and what we know what kind of representations exist in the brain and how brains do those things, and what kind of architectures exist in the brain. And I'm kind of excited to see that kind of, sort of,

33:08

sort of things coming together so that we can make sense of how complex behaviour is sort of things coming together so that we can make sense of how complex behaviour is produced by brains. Yeah. Okay. So it already has impacted in a positive way, my life persona, that before I told my wife isn't native English speaker and I always said was, when I had to write an important document to proofread for me, now I have GDP.

33:31

To do that, my wife seems very happy to don't need to do that anymore for me. But jokes aside is exciting. And the problem is, what's the goal here? And I think we need to keep separating the goal quite well. So my goal is understanding the human brain works and things. And it's not like when planes start to go faster than birth, the field about any technology as is finished, right? Is it really, and I think most of the problem comes from the confusion of the objective that we have.

34:04

What, as a neuroscientist that is interested in AI, I'm actually most interested in even when my current work, I'm looking at it, I don't know if it's going to be close, I don't know if about never trust is going to be in 70 or 80 years. How often joke with them is that those seven years keep moving is like is that seven year is their idea of time if you want something that seems far but not too far, but far enough that people forget so.

34:36

So if you don't want to do something, say I'll do in seven years. So that seven is the magic number. So the things that I'm most, I didn't say at the beginning, what actually do I study, value based decision making and is the decision that, for example, you've done a very strange value based decision making today, sitting down here in a dark room, listening

35:00

Listening as rather than either no going to the cinema or doing something else. And I'm curious why you did that. Now the thing says that unlike machine, at the end of tonight, you're not going to get point for the decision you made, right? So you made this decision, you made the judgement call to do that rather than going to a restaurant or watching a movie. And he's not that in cash Well, at the end was parallel, you had 1000 point by doing

35:28

the things though instead machine, the way in which we train is exactly the opposite, that they always need to have a point at the end of the things that are being traded for. And my big, big question is, would we get to a point in which you construct your own value while in machine at the moment, even the most sophisticated, we always give these very artificial exogenous value, you want the match, you get one point you get traded, but you never get point in life, you build to the point in your head, if you if you do probably don't even do that, you know, tonight, the decision has made you have you made a good the right decision is not going to hands up with a score of how you how many points you made, now means that you constructed the value of reality around the view. Now the question is machine are really far away from that. And if you want to make them be more like flex, while acumen there might need to start to do that. But that's becomes a problem as well. Because Can you imagine machine startup preference, liking and disliking things?

36:46

It could start to get, you know, humans have a very strong preference. And according to their preference, they make very important action. The so that's that's a big question that interests me, I'm still not very much idea what will be the architecture, but also worries me a tiny bit. So the intersection with morals essentially, I mean, if you like ice cream, and you dislike Madame Bovary anything in between, right. And in the you know, you, you can a people that cold Caswell, I mean, it's like,

37:24

like, when when you build your own value,

37:30

it's the only things you can do because we live in a world without tags, right? And we, but then, and we haven't even started to approach this problem. But sooner or later, we should, and probably we will, because because the barrier will be heavy than that, as he was saying, the training finish you you can train yourself, because you can actually build your own value. And I'm curious how you do it? And sure, we'll let the machine do and will they do? Well, I don't know. Maybe I'm just a psychopath myself and

38:10

interesting question.

38:12

I guess, in the psychiatry realm, I mean, the bit, I'm the bits I'm most interested in is really understanding, you know, mental health disorders from a computational kind of perspective, because, because

38:30

I think

38:32

it just is a much richer, deeper understanding of what they're about than any kind of very biological reductionist type of understanding because it because when, when, when you view people as having a model of the world,

38:47

you can't avoid all of their background, all of their childhood or their environment, all of their social influences, and, and the biology that forms it. And it's not just the, you know, it's not just the serotonin and whatever, you know, it's the whole thing. And so, understanding that is, is super interesting, but I think in the next 15 years, it'll take quite some time to get that I think, in the immediate future.

39:15

The most progress is going to come from these kind of black box machine learning type approaches that just say, give this person this antidepressant or this person, this antidepressant, and it's not totally clear why but it seems to work.

39:28

But even that will be great for now.

39:32

I think there's like I should have told you, they're gonna we're gonna get the audience to rate you afterwards, there is going to

39:39

so it seems like actually, what you've all done to a certain extent is draw a distinction between applying applying these machine learning models as a tool. So you know, diagnosing a thing or deciding whether that's a label of a cat or a snake. And I think probably you'd all agree that's already being useful in neurosciences.

40:00

Don't be useful in the clinic. My lab use things like deep lab coat, which is a Bayesian automatic way of labelling animals and boxes and things like that. And I think that's unambiguous. What's, what's kind of interesting is the way you've talked about using using machine learning models to learn things about the brain, not just to like sort out your data. But you know, like, can I? Can I look at how the machine learning models solve this problem? Or how it compresses data in RNN or a transformer, and then make some inference about how the brain does that. And I think you sort of came closest to saying that picture is, is that true? Do you agree with that distinction? And are you optimistic about the second half? Yeah, so I think the first half, there's like no, no questions that it has been really useful, it has really helped us the noise, the data is how I would say it or label it. And that's been incredibly useful. Everybody uses it,

40:53

I think. And the second half is like using these models as models of the brain themselves. And I think there's appeal to that. Because up until now, up until that transformer architecture became popular, they had lots of inspirations, they have lots of properties that exist in the brain. So they were very useful. And especially like sort of drawing is an example from my research, I found that sort of two brain areas, which had very similar representations, but had very different responses when you perturb them. And this requires like a multi region, sort of

41:30

network which has multiple different neurons, so that they can be robustness and that network. But also, you can put up specific projections from one brain to the other. So it requires this kind of architecture to be able to study it. So for my particular research, having these tools available to train such kind of big RNNs was really useful because I could make this model, which replicated all the properties that are found empirically. And then I could study I had like a fully observable system, in which I could do all kinds of experiments and understand how it works. And if there are any normative advantages to having this computation be performed this way. So for me, it was like a really useful tool to advance. In this case, it's both a model of the brain, but also a tool to study the neural processes. So yeah. And actually, I was really struck by something you said there, which reminded me strangely a review or comment I once had, but it was

42:22

it was this right. But

42:25

as you said, you had a model that you could sort of was fully observable. And and there's often this sort of saying that he will say, Well, why do you want a modelled brain with a deep network is like another black box. But I think you'd agree base what you just said that it's totally not, it's, if that's a black box, may if as a black box, you've never tried to study the real brain, like, at least with a deep network, you can, you know, you can look at what it's doing any point in time, you can wind it backwards and forwards. When you get complex enough, you need to be always there's a story I love to say that is a study from this Argentina bride that it was called Borg. Yes, that was the cartographer of the empire. And he says they existed an empire in which the art of cartography become so sophisticated that the map were more and more precise, until 1.1, cartographer made the map of the empire that was as big as the empire in which every point on the map correspond to a point on Empire. But they quickly realised that that was completely useless because it was the largest empire. And since then, the act of cartography died.

43:30

But the thing is, is you're right. First, I think many people will be surprised to the fact that what she just said that we actually do experiment on this machine on this. So we build the network, but because we haven't built into it what it does, then we need to test it almost like relation, the network and things and I agree with you, is simpler doesn't scream when you do things. And the thing says, though, that will, we will be apt to be very, and there's almost like a philosophical epistemological question. What if interesting things started, only a raise a certain degree of complexity. Well, back on square one, we are now on a map. That is the sides of the empire. And a minor.

44:25

Yeah, and the good things about yeah, sorry, I'm I think the gap you're talking about is that once we have an RNN, we still can't fully understand. It's still kind of a black box and the way we don't understand exactly how it works, exactly what it's doing. And that so there's a gap in having the tools to be able to understand that. But from where I started, I think developing those tools is a prerequisite for understanding the brain. So I think it's almost pushing us in the right direction of like, if you have this kind of network of balls, how do you study them? How do you back engineer what's going on?

45:00

And this is helping us move in that direction. And hopefully it will come back to just doing that with the brain data that we collect, and we collect more and more of it. So to me, it's a kind of like a useful detour. So to say of like one kind of class of models, citizens buying the right kind of methods that we can use. And I agree with you. Yeah.

45:17

Is a grey box.

45:21

Golden black?

45:24

Rick, do you have anything to add? No.

45:33

So I'm aware with quick actually times going faster than I thought on this, this first off um, so I want to get to I want to ask you maybe most might be our last big science question before we switch to the the interval and then on to the social aspect.

45:47

What are the big? What's the big like, could each of you just take well, you've got a minute each basically, what's the starting over there? What's the sort of next big question? In your field? Related to neuro AI?

46:05

Medium sized question. Yeah.

46:09

I mean,

46:11

I mean, I think it's going to be

46:14

people are starting to find these, these mysterious signatures that seem to split patients up into groups that might respond to an antidepressant, one kind or another. And then the big question is, what's the mechanism? And so

46:35

answering those questions means having the traditional approaches like animal models, and do and being able to perform causal interventions in those animal models, but with all of but trying to do it in such a translational way that you can marry up the the machine learning results, the humans and the animal models. So there's even more kind of complicated to do get them all in line. But I think that is the year to get to discover the mechanisms behind these decisions would be would be most useful. Because presumably that way more treatments lie as well.

47:11

I think I don't know if this is one question. But I think it would be like given a model and our internal facade, can you figure out exactly how it's doing what it's doing? So given something like a weight matrix, some architecture basically, which is embedded in weight matrix, can we can we sort of glean from that what function is implementing How is implementing the dynamics, I think would lead to big advances in neuroscience is the question to answer. Yeah.

47:41

One I mentioned work, whatever you can from that, I think is something interests me a lot. I would say another one is

47:50

how we learned the right prior. I know it sounds very cryptical what I'm saying, but most of machine learning has been trying to avoid doing prior information, let told the data learn. And that comes to the problem. We discussed it tonight. If you want to start with zero prior, you need a lot of data. Human Use prior and you have a prior information. There is a famous case of one Tesla that crushed in, in in a car, and because of that detecting system didn't see the car anymore. And you would say actually that their camera was better than the AI. So a human wouldn't have seen it. But human would have known the object did not disappear. So human eye that prior in the car didn't have that prior

48:40

how we learn the right prior because prior are really useful. If are the right one. If you learn the wrong prior, you're kind of screwed, or better, worse, they starve to death.

48:52

Then you're really the big question is how we human. How we learn the right prior? that I think would be Did I explain enough World War means? Learn your prior well in life?

49:10

To your kids, like, kiddo, learn to your right price and putting in the Word. There we go some sound advice in

49:19

this session, and thank you ever so much to our panellists. That's been an amazing discussion. We will now have a short interval about 20 minutes

49:38

Welcome back, everyone, first of all, and welcome back to our panel. And we will start with some questions audience questions thanks for for filling these out. And I unless it says otherwise, I will read out the question and invite each of the panel members to comment with their answer. So will we

50:00

Get a AGI defined as an AI that can do everything Caswell does Wait, someone wants to make you redundant. before.

50:11

Before we get close to answering the hard disk questions in neuroscience, I really like likely I think we'll have very sophisticated automation systems, we already have them. And I think they're gonna get more sophisticated very quickly. I don't know, they'll replace castle, but they can do a lot of folly will be able to do a lot of things. I think it's probably going to be longer due to figuring out exactly how

50:34

our brain and how human brains work, which is like an even longer road. Yeah.

50:41

Next one, are you optimistic about the hierarchical predictive coding theory, exploring the neurobiological basis of autism and schizophrenia in the future? That sounds like what you're up to? Yeah.

51:00

Yeah, so I think I think I, this is something I've worked on in my PhD and then a bit subsequently.

51:09

For people who don't know what the question is about, it's basically this.

51:18

One thing that we think the brain does is this

51:23

process called Bayesian inference, where you use some priors use some experience that you already have to interpret new data coming in. And this is much more efficient than trying to figure out just purely from new data coming in what's going on all the time. And this hierarchical predictive coding is one kind of network that can do this, essentially.

51:45

And so the idea was that this network, if it was imbalanced in some way, it might go wrong. And

51:54

it might lead to the symptom, the symptoms that we see in schizophrenia, and also in autism.

52:02

So I think I think it does go some way to explaining some of the symptoms, for sure. But I think

52:11

there's a lot more that the brain does that we that you can't fit into a simple kind of predictive coding scheme. So, so this, this predictive coding hierarchy is a good way of explaining the kind of perceptual bits of the brain.

52:27

But then the decision making part and the more sophisticated kind of memory parts. So everything involving hippocampus, prefrontal cortex, is probably some doing something very different. And I think we need very different models to think about those. That makes sense. Thank you. And I love this question. And it's one that I might actually be able to comment on, which makes me happy I know something about what's happening in neuro AI. So are you aware of any machine learning approach has been developed that can predict the effects of specific drugs at the preclinical stage? So for example, can we use AI and machine learning to accelerate the transition of drugs into clinical testing, but also lower the risk of adverse effects? And how close do you add? Do you how close do you think we are to the development or realisation of something like this? I guess this one.

53:29

Sorry.

53:32

So yes, I do. People are trying to do this. It's definitely not my area of expertise. But I've seen a talk somewhere, I can't even remember who or where it was, but but people are trying to do exactly this in order to.

53:49

Yeah, they're trying. So what they're trying to do is, is

53:55

enter all of the compounds that are known that have since synthesised by all different pharmaceutical companies, and then enter all of the different paradigms and chemical properties and tests that have been done on all of these compounds, because there's maybe 100, different tests are 100 turned to different tests, but only probably 10 or 20 are done on each compound. And so if you want to then predict the outcome of a test on a compound that hasn't been done on that compound, compound, but has been done on ones with a similar shape, you might be able to do that. But obviously, this is a really complicated thing, because you've got hundreds of variables on one side, and you've got these complex structures that interact in different ways on the other. So yes, that tried to do exactly that with drug discovery to speed it up and to reuse. Repurpose existing drugs. Yeah. And maybe I can come in quickly because it gives me an excuse to plug what will be our next pre recorded episode, which is that this morning Steve, who's our other co hosts, sadly, not here today. In

55:00

interviewed Sonia Gandhi at the Crick Institute. Now she is working on Parkinson's disease. And they have used methods to grow patient cells in the dish, and then use AI and machine learning to predict whether the particular properties of a cell are linked to that person's clinical symptoms. So although it's not exactly the same, I can kind of see in the long term, there could be an option to combine these approaches and say, Well, look, this this person cells tells us they're likely to develop the disease at this age and have these symptoms, and therefore this person is more likely to be respondent to these particular therapies. So good. I feel like I've managed to contribute one thing about neuro AI today. So I will ask this question, and then I'll hand over to Caswell for the for the rest, how important I think this one might be for you. Deeksha. How important are the parameters of an artificial neural network? Is it important for neuroscientists to study and refine the parameters in order to improve the performance or the outcomes of the artificial neural network? They're incredibly important, like, there's definitely like some sloppy regimes in which the function output you're looking for will be present. But they can't be. I think the problem is that those parameters don't tend to be continuous that you can get good behaviour in this very distinct set of parameters. And then also in this very other sort of space, but you don't expect it to produce that output. So that produces special kinds of challenges, because you're getting very different solutions to the same problem.

56:36

And it's incredibly important to like, actually, look at what kind of parameters you find the

And it's incredibly important to like, actually, look at what kind of parameters you find the solutions you're looking for. And then try to study if those correspond to what we expect in the brain, if they like,

56:49

sort of match the expectations we have. It's kind of hard to answer this question and like a general way, but yes, parameters are important, and we should be looking at them. Yeah. I totally agree whether we're this is one of my massive bugbears.

57:04

Obviously, not all networks behave like the brain. And if we can understand the set of like the parameter space where they do behave like the brain, then we've just learned something really important about the parameters that applies, right. Anyway. But I should read out the questions. I totally agree. So I have two really good question. So do you think we have a research sunk costs with AI for neuroscience? Are we focusing too much on drawing parallels between the brain and AI at the cost of exploring new and different theories of what the brain is doing?

57:35

There was something said that at the time, that I think Rumsfeld said it doesn't sound funny intellectual to quote, he said, there are known unknown and unknown unknown. And the most interesting are the unknown, unknown. And when you say does an excuse to invade Iraq, and there wasn't really like, but there is something

57:57

profoundly philosophical into that, that we do not know where things are coming from because we, the unknown, unknown, you don't We don't even know that we don't know. And maybe, that the answer to that question is,

58:14

we don't, it could be that this might be completely red herring a wrong framework. And somehow the right framework will arrive, but it's very difficult to answer. Because it's going to be probably an unknown unknown, or that we don't know that we, we should know something was good.

58:35

So this our last written down question, but there will be if anyone's got any questions that were asked after this, there's gonna be opportunity.

58:42

So this question starts. I don't know anything about machine learning, but decision making is adapted to your experiences. And I think around the world, different cultures and values teach different forms of decision making decision. If these models are developed with a western point of view, are we in a sense, ignoring or marginalising other populations? Ie are we reinforcing a single point of view in areas that can't relate to these decisions where this reasoning isn't applied? Anybody wants to take that through

59:12

and there are studies that show that there is an there is a real problem even just in the training set we're using is like we're using English speaking internet and even face and thing there is a huge problem in fact, one of the biggest issue there the charge ATP or these other things, sometimes seems very intelligent, but sometimes they start to become really cold and answering you and the reason is because these these these, these models have been traded on racist stuff and thing and now this company in order to don't embarrass themselves, had somebody in engineering. Actually what they do is really strange and I know that because different reason I was in an ethics board and

1:00:00

The so you you actually have people in the in other country that they they have to mark if the answer is offensive and then they can't craft out of the model. And that is reason because of the training. And there is a lot of moral issue there as well. Because also, what is offensive in a country might be different than other things. Plus there is this new colonials paying that date, those work is being done in country where it's cheap labour. So super problem somebody else want to add? Yeah.

1:00:35

Yeah.

1:00:37

And, and the something we need to deal with is very strange. You'll be it'll be surprised, trying to get one of these journalists or model to tell you something's uncomfortable. It will refuse not because he can't do it, he can do it and probably will do in a very biassed offensive racist way, has been handcrafted out that thinks, and that's a big issue. Yeah.

1:01:04

So I wanted to add to that, in the sense that when you're looking at the clinic, and you have a system that's making a decision, I don't know anything about this, though sort of stuff. But I imagine like, it's been trained to be unbiased. However, different populations, kind of normally have doctors that know the reality of that population. And so when you're making a diagnosis, I'm not a doctor, I imagine the living situation of a person is taken into account and the kinds of therapies that they can do, and the kind of things that you prescribe to them. So if you're teaching it without barriers, then isn't really applicable to situation which people do have barriers to what they can add. Can I just repeat the summarise the question just for the recording? I think it's great point. So it's around use of AI in diagnosis, and actually, the fact that people from different populations may have barriers to what they can access. So how do we incorporate that into a system, which we know has bias? Is that a good?

1:02:14

It was more related to that if I Yeah, there was just saying like, you know, a doctor, the lives of experience in their population may know, things that these AI system that have been trained at the London, they do not have, and we are very brashley, ignoring that our at our own peril. Is that complete your question? Yeah. So I've done so much.

1:02:45

So there is so there's a very pertinent example of that, but the potential for that in

1:02:52

and schizophrenia, psychosis world which I live in, and so,

1:02:59

so black British people are more likely to be diagnosed with schizophrenia, psychosis, if they present the same symptoms that that the white British people do. And they're more likely to be sections like, so put on a ward against their will, and various other things. And,

1:03:20

and, obviously, if all of that information goes into your training data, which is all the training data that you're going to use, then that training data will not be unbiased. You know, it won't, it will be, it will be the opposite of unbiased, it will, it will be much more likely to diagnose schizophrenia, or psychosis, or whatever if it sees a black person than a than a white person.

1:03:46

And so,

1:03:48

and some of these, some of these biases we're aware of, some of them were not. And so some of them, you could try and tune out in this way that Magneto just described, but there may be many that you might not. And so, yeah, it is a it's a very real problem. But to reassure you on this count, I mean, on this particular account, I think we're very unlikely to be using these kinds of methods to diagnose psychiatric problems anytime soon.

1:04:22

Because there's so many other problems, this is just one of the millions of problems.

1:04:27

But you know, there are, there are scenarios where things think that you know, legal scenarios where potentially legal processes can be potentially replaced by these AI

1:04:43

machines where you definitely don't want these biases to creep in. And lots of lots of, you know, credit scores, credit checks, all these kinds of automated processes may

1:05:00

decisions that are unaccountable and may be due to biases in the training set. So this stuff is a very, very live issue.

1:05:07

Totally agree.

1:05:09

A question down here, I suppose it's kind of links in with, with your question, you briefly touched upon how AI models can be used in areas where, for example, there might not be the same amount of doctors available.

1:05:22

But I feel like this creates another problem. Seeing as models have to be trained based on data that you get from a certain area, if you want to do it unbiased or based on that area. But if there's not the right amount of state want to use it. So there's not enough doctors to give you that data, then how would you train? Suppose it kind of feeds into the same problem? But how would you go into solving that? I'm going to repeat that quickly for for the record. So the point is, basically, because we've already got this sort of resource disparity between place both in terms of doctors and other resources, how are we just going to compound that basically, by training, training these models like, you know, if there's a region hasn't got many DBS? Whatever, no training data, therefore, they're underrepresented, therefore, the cycle continues?

1:06:09

How do we get out? I mean, I can give a very quick answer. So so there's this field called transfer learning, which is really all about how we can, how we can keep what we want to what we want, and discard what we don't want in new scenarios. And,

1:06:30

and, and also, funders, like welcome are realising that this issue is of massive priority, and they're so and so they're funding much more research in resource poor areas, primary research in resource poor areas, to provide these kinds of datasets that will be useful. So I think those dual approaches would be very helpful.

1:06:51

Great, if we got any, there's a question there.

1:06:58

What sort of impact your AI has on the world?

1:07:04

I'm going to repeat again. But what impact does neuro aI have on the world of dementia, we're surrounded by experts. So this one

1:07:12

who'd like to say, you're gonna get their time, Rick.

1:07:17

I mean,

1:07:18

anyone else have it?

1:07:21

I can eat. It's not really my dementia is my area. But I'm more of a cell biologist. So I can say that we are putting in grants which are so far unfunded. But eventually, we would like to use machine learning and AI to see what we can't see. So we are doing a lot of cell biology, we are growing neurons from patients who have dementia or don't have dementia. And we look for specific things in the cells that we know are associated with the disease. But it means that we are probably missing things that we don't know what to look for, or they are too subtle for us to notice by. And so what we would like to do is take large datasets that we're accumulating through gathering images and images and images, you know, my people in my group, one of my groups here, I can see you nodding away spend hours on the microscope. So we develop these really rich banks of image data. And we think, Well, if we can ask,

1:08:20

via machine learning, can you cluster these into which other dementia cells and which are not the dementia cells that might allow us to detect changes that are happening earlier in the disease? And I think one thing we all agree on in the dementia world is that if we are to develop new disease modifying therapies for dementia, we need to be intervene as early as possible. So actually, those early changes that we might not be able to detect, but really powerful microscopy may be able to will be really valuable. I guess the other area where I could see utility, which is not not my area at all, but something that maybe Rick can comment on, is around things like analysing MRI scans and things like that. So we can see, you know, we have been the royal we I haven't done any of this, but in the field.

1:09:13

Emily, Emily's

1:09:15

in the field, you know, there's some really nice studies that have followed the natural history of dementia in people who carry genetic mutations. And these are the only individuals where we can really say, we know that you will get dementia and when you do get dementia, we know what's causing it. And what we can see is that there's huge kinds of structural changes to the brain decade before somebody develops symptoms. Now whether we can actually expand that using machine learning and AI to look in the general population rather than these really rare genetic families. And whether we can use that to even move earlier. To see things happen in you know, before again before we can kind of detect it with it.

1:10:00

Non machine learning methods is something that I think is really exciting.

1:10:05

So yeah, not quite my area, but that that they're the two kind of themes where I see the most potential if you like, Rick, I don't know, if you want to correct. I'm happy to be corrected.

1:10:17

No, no. I mean,

1:10:19

the interesting thing that might makes me think of is,

1:10:24

you know, it's

1:10:27

would you want to know, if you were gonna get outside was in 20 years, but you could it's an engineering challenge. But the clinical use

1:10:37

funds another question. Oh, yeah. Yeah. Yeah. Once the, if there's a disease offering treatment, for sure you do. But yeah, there's often there's, there's often just, there's often a divergence between the engineering challenge and kind of pragmatic.

1:10:54

I 100%. Agree. And I think, you know, we, hopefully our next live event will be on Alzheimer's disease, I've just committed as to it as cause well looks at me in horror.

1:11:05

No, I think it's one of the things that we are seeing is we're starting to see success in clinical trials. And what I mean by successes, we're seeing disease modification, so we can slow the rate of decline in someone with Alzheimer's. But we need to know who to give those therapies to and the earlier we give them, the more successful they will be. But actually, I agree with you do we want to be telling people Oh, yeah, we can see these changes in your brain, but we can't yet give you these treatments because they're not approved. So there is a kind of balance there. I completely agree.

1:11:44

question at the back there.

1:11:47

So I was thinking about what you said before about kind of metacognition. And the question is kind of what we think about metacognition, evaluating your own positions and thinking about whether that was the right judgement, why you went to that. And the problem with machine learning models is that they can kind of understand why they did a specific decision. But oftentimes, when we think about metacognition, it's so intertwined with consciousness. So when you think about the whole metacognition in machine learning, can you separate that from the whole consciousness debate for the day? I

1:12:20

guess Lanford Can you summarise the question before you answer it, please. First of all I need to summarise I wonderfully managed to pronounce my name, I normally get told Benedito and things your benefit was fantastic. And

1:12:38

you say that a very question is very dear to me, as well as the problem is I spoke about metacognition and the fact that that, you know, this architecture may miss metacognition and maybe maybe useful to have it for all the other things. But then there is a very strong link between metacognition I mean, it's unfortunately we don't have Steve Fleming here that is Mr. Allen metacognition people like Steve and other people have been thinking a lot about the relationship between metacognition and consciousness. And so the question the natural question is if we're going to have eventually system we metacognition then the question is, how are we going to make conscious being on the way? And

1:13:26

it's very interesting question.

1:13:30

There is a big reason I always see the difference between the way I studied metacognition I've been coming from economics, I was much more practical in the study of metacognition

1:13:42

compared with people like Steve, they studied in their own sake, and from because, you know, in economics, there is always you have to convince economists, they are very smart and very stubborn. Why you want to study metacognition? And, and, and this, I was always trying to find the reason why metacognition is for and metacognition can be for maybe correct because you know, it comes after you made a decision, that's the strange things about metacognition is almost too late to correct whatever you're done. But the fact is, we don't do things only once in like, we repeat them. And now the thing is, is the fact that the complex system as a is almost like, what's the point of IB consciousness, you know, philosophers have debated this thing. It's not like evolution likes to make conscious thinks there might be there is a bar there is a point in which to have this sophisticated level of control. You need a second order system, that is what metacognition is, and then consciousness is going to become a byproduct of that. And some people that work in AI, the I know they it all, you know if consciousness is important is gonna erase on the way to it

1:15:00

I feel that there are very good reason why, you know if evolution evolution has endowed us with consciousness because seeing Right, right, this is an interesting experience. So it might have been very strong pressure for it. And we're going to face the same pressure in those artificial agent and I think we are already facing it because one of the point I've made in some article that the reason why this human can learn with little data you know, the things I've been keeping telling tonight, learning with little data wanderings really helps to learn where little data is having the second order system.

1:15:42

And eventually one system alone with little data. So eventually, we're going to face the fact that we need to this second order system, and then is going to just, you know, generate protocol, viciousness. Maybe, maybe wait when steps come back, and they will finish to answer the talking about Steve. CV is almost like

1:16:06

my best friend I just like a matter, Steve, if you're listening me now on the podcast. We don't want Steve to think that we can't cope without him, though. So as well as doing just fine.

1:16:20

Any other audience questions like Caswell with hair blow?

1:16:26

Question in the back there? Hi, it has been thrown up quite a few times today that

Question in the back there? Hi, it has been thrown up quite a few times today that

1:16:33

AI and kind of what we can do with AI. And what happens is that learning can like what humans can do is generalise learning very well. And as you guys said, a child can only like can see three or four horses and knows that anything that looks like that's a horse, whereas for AI, it needs much, much larger datasets for that. Are there any psychiatric conditions where generalising is impaired? And can the AI help us particularly understand these?

1:17:04

Are you happy to summarise the question? So we've Yeah, we've talked a lot about generalisation of learning. And so the question is, are there any psychiatric conditions where this kind of generalisation is impaired?

1:17:19

Um,

1:17:22

so

1:17:26

maybe not. We really need generalisation. I mean, well, in some, yeah. So in some conditions, there's overgeneralization because there's too much not too little. So, for example, in post traumatic stress disorder,

1:17:41

a really unpleasant experience, say a red car crashes into you when you're little

1:17:49

means that subsequently when you see other red cars or you're near a roundabout, you get the same panic response that you had back then.

1:17:59

And that's an overgeneralization response.

1:18:03

Rather than under generalisation, one, there are

1:18:07

there are things that the hippocampus does, that definitely are not there on or severely impaired in schizophrenia and psychosis. So for example, I don't know about general, I don't know if anyone has ever tested

1:18:25

generalisation itself properly, that would be an interesting experiment to do is to broader concepts, testing one experiment, but but they've they've tried coupling, different inferences together. So learning that A and B go together, and B and C go together. And if you learn those things, then your your hippocampus also learns the ANC will go together, even though it hasn't seen those things together.

1:18:53

Which is what I think your entire projects has

1:18:59

been for patients with psychosis, schizophrenia, can't do that. They're really bad at that. And we do know that, but that's not the main problem for them. But it is it is an indication that their hippocampus is not is not working properly. So I imagine there will be some generalisation problems. Well, I didn't do it. Yeah, I was gonna add that some of the generalisation problems that I was thinking of when talking about it are like, even more bizarre than what we see in like human psychiatric conditions. So if you change a pixel neural network will detect it's a totally different thing. Even if it's looking at a rat. What very clearly is a rat. So it's like very brittle sort of failure of generalisation that I was taking off under these more complex, interesting scenarios, which I don't know if somebody's looking at that. I'm sure they are. But yeah.

1:19:45

Super, thank you. Well, we've got about 10 minutes left. And thank you so much for all your questions. We were having a slight panic before that. There will be no questions and now I can see this.

1:19:57

But it's brilliant to have such engagement.

1:20:00

because we're coming to the end, I really don't want us to miss the opportunity to hear a little bit about the career journeys of our panellists. And I think this is the other wonderful thing about brain stories, is what we've heard so far from our previous episodes is, everyone takes a different path to arrive in, in their research area. So maybe we can just spend a couple of moments each maybe I'll start with you, Ben detto, talking about your career journey, how come you first became interested in the brain, and what brought you to where you are now. He's actually when I was at the beginning of beginning a few years ago at UCLA, I did a stand up comedy about my journey.

1:20:43

He was so if you want to know really about the journey with all the funny bait is that most of the things after experiences some for comedic effect, exaggerated,

1:20:55

there is strange, this is the beauty of neuroscience, you can reach here from very, very strange paths in I started to some my mom at the pharmacy. And in Italy, it was almost like almost because there is I unemployment in my area was almost like, Oh, you're lucky, you're gonna have a job as a pharmacist. And that things I was dreading these things a lot. And then I by really liked chemistry and things I tried to do, the only degree I could do that didn't allow me to go into pharmacy. It was a very convoluted plan,

1:21:32

studying the same things, but with a degree that didn't allow me to be stuck in, in a little pharmacy in South of Italy. And then I start to get really into the molecular biology, developmental biology. And I came here to do a PhD in a completely naive way. I it was a Wellcome Trust PhD, but I didn't even know the welcome. You don't write with. You've write with to hell. So I thought it was just a very welcoming programme for just arrived. The thinking was UCL welcome view.

1:22:09

I pretty sure that they took me because my English was so poor, that they thought this guy must be a genius.

1:22:16

To study English. In during the interview, I even say there was this protein was called FGF. For essay, have Jeff Quatre. And people were looking at me sounds bad.

essay, have Jeff Quatro. And people were looking at me sounds bad.

1:22:31

I mean, to shorten the story, but then this programme forced the welcoming programme, forcing you

1:22:40

to do things weren't out of your comfort zone, because I was sure I was coming here to do molecular biology, but they forced it to you to do something different. And that's a great thing, because it forced me to do something that I wouldn't have done. That was neuro imaging. And I went there, mostly because the building look nice. And then actually, I actually thought he was actually fine. And then in the in the PhD, again, there's so many points in life in my life by thinking leftover, things just go in a very, very strange way. I was working on attention, nothing was really working. Herding was really not going well. And then I watched the movie like beautiful mind. And, and I liked it a lot. I read the biography. Somehow I stumbled on the I was searching about John Nash on the website. And they actually stumbled on Kahneman and the framing effect. I didn't know what you were, then I come up with an experiment. Along the line six months later, I get Kahneman, emailing me and say I've read your study on the framing. I want to meet your lawn. And so things could have gone in a completely different way. And so embrace the cows. That's my suggestion. Don't plan too much. And, you know, maybe I've been lucky. Sometimes. Sometimes things go without your control, trying to do what you find fun. And be able, maybe this is long advice. I wanted to become a poor poor when I was little, so I'm guessing giving advice.

1:24:19

Because my mom told me, I asked my mom, where's the pope keep his wallet. And she said, the pope doesn't have a wallet and I say how does he pay for it? I mean, well, so pope is on television every day. And they say there are other people pay for him. And I thought this is a great job.

1:24:39

You don't have a wallet and everybody pays for you.

1:24:43

So anyway, then I discovered to become a puppy or become a priest and become boring gold. So I like the only I shouldn't become in the middle age when you just become Pope. Anyway, my only advice is, be honest with yourself. There will be some things you

1:25:00

might be good at it. And some people are not very good at it. So really wanted to do molecular biology. But I'm terrible with my hands. I'm the clumsiest person in the world. And I was fighting against it for a long time. Sometimes you just need to accept, and you just like, can't don't think as a sunk costs, you spend money on that, and things, just maybe you need to change. And that might be something I would have been, you know, pretty terrible with my very poor motor skill. And then I discovered doing computational modelling, you can have really poor actually, compared me, my computational colleague, I'm actually

1:25:37

sophisticated and well

1:25:40

have passed the Deeksha any aspirations to be the pope when you were a child.

1:25:48

So I grew up in Delhi, in India, and I was always very interested in biology, I didn't want to be a medical doctor. So if like your interest in biology, parents tell you, you've got to be a doctor. And I didn't want to be that at all. So I just did what everybody else was doing around me was just like prepare for this engineering exam. So I don't know if you know this, like the six engineering colleges in India, which are like highly reputable. And anybody who wants to get a job just tries to get into those universities. So I was preparing for that. And I got in your engineering exam and studied engineering undergrad. But in the back of all of this,

1:26:25

I had this like, sort of when biking around in the city, I had this like, sort of awareness of the limits of my sort of say consciousness and freewill in which like, I would avoid a pothole. But I would know that I didn't see that pothole, even before I sort of averted and like kept riding. So I was very interested in how that happens. So I was reading some philosophy at the time, through which I made my way into like pop neuroscience. And then in undergrad, I met a group of people who were interested in neuroscience and be like kind of small group called Science coffeehouse, they would chat about these things. And so I developed this paddle interest while studying my regular engineering things. And then in my final year, I got to study with this professor who was actually studying robotics, but he was interested in how robots could learn language and learn from interactions in the world. So that gave me like, more of an insight into how we can study these things about How do humans develop, learn these things. So I applied at that point, I was very keen on interacting with real brains. So I applied for like PhD programmes, and that got lucky and got into one and then worked with actual Jordan brains recorded with them, sort of like part of their study them. And now I'm here. Yeah, with by you sort of, in a way very straightforward part. But for a while it looked like there was no part for me to get into that. Yeah.

1:27:51

Yeah, so

1:27:53

yeah. So a school. Yeah, I had no idea. I had no idea what I wanted to do. I didn't even know if I wanted to do science, sciences or, or kind of English and stuff. I chose sciences. And then I kind of went, I kind of copied my friends instead of ignoring my friends. I went and did medicine. And then, and then starting to do that I got more and more interested in the brain because I just found it more interesting than all the other organs. And I thought I wanted to do neurology at first. But then when we actually did, and so I read a lot of kind of neurology books. I've read lots of Oliver Sacks, books and any of your work for me a lot of sacks, but they're fantastic books, you should read them. But then when I met some neurologists and I met some psychiatrists, I discovered that Oliver Sacks thought he was a neurologist, but he was actually a psychiatrist.

1:28:46

And a neurologist body.

1:28:50

Like a lot of your old colleagues.

1:28:53

And, and really, I did a psychiatry placement with

1:29:01

a fantastic mentor, this woman, Mary Robertson, who used to work with she was an expert in Tourette Syndrome who used to work at the National Hospital. And she was just super inspiring and very, very encouraging, and,

1:29:17

and really just completely captivated my interest all of a sudden, and made me and I met people with delusions on the ward and hallucinations. And I just found the one that began the Pope. You right, yeah, exactly. Yeah. They were the Pope.

1:29:34

I just found it was. So I just thought it was such an interesting question about how these things arose. And could we do any better in trying to treat them than we're doing at the moment? And so ever since then, I've been super interested in that but then, but then I didn't actually do any research for a long time I qualified and I'd worked as a doctor for seven, eight years.

1:29:58

And I kind of knew I wanted to do research, but

1:30:00

I didn't really know which group or what to do, and I just kind of bite it waited and waited. And then I ended up emailing various. I did this MSc and philosophy of mental disorder and stuff like kings.

1:30:19

And then, and then I read some of conferences and stuff who, who works at UCL. And that I found super interesting. And I emailed him, I emailed a bunch of other people around neuroscience kind of investigators around London, just asking them, do you have any places and and to my astonishment, loads of them said, yes, come meet me, like, a medical student emailing Consultants is like scum of the earth. They're not.

1:30:52

You know, they wouldn't give you the, you know, the last breath, but the but but PIs are generally keen to hear from new blood, because you might think, yeah, exactly. Yeah. Yeah. cheap labour.

1:31:08

Yeah, the value function is different. So then I Yeah. And so then, yeah, that's how I got into the that PhD. And then it's just yeah, it's taken off ever since. But there was no grand plan. It was all just following things that were interesting. So yeah, my advice would be to find things that really you find interesting and motivating. So expose yourself to as much different stuff as you can. And then when you find something like that, that you like, don't hold back about getting in touch with people and, and putting yourself out there and finding out what what it's like to do that.

1:31:46

So we're almost exactly out of time. But I'm brainstorming is we always end with the same question. I'm gonna ask it to each of you. In turn.

1:31:57

You're gonna start out Listen, no.

1:32:01

So what's your favourite fact about the brain, Rick? So, yeah, I don't know about one fact. But one very exciting thing, which I read two weeks ago, which is the most exciting fact of the moment was, there's a researcher called Helen Mayberg in New York who have for 25 years has been putting deep brain stimulating electrodes in people's brains to try and cure them of intractable depression. And we've mixed success over the years. And

1:32:34

she's just published in Nature, this study showing that they've applied some machine learning methods to these, the readouts from these electrodes, because you can record as well as stimulate. And they showed that you can predict the onset of depression one or two weeks in advance in these patients from this EEG readout.

1:32:56

From relatively simple

1:33:00

properties of the data, and that is pretty exciting and cool, and potentially quite something. So that is my most exciting facts. Those are pretty good facts.

1:33:12

Deeksha. So my fac the preface, my fact is that brain has 86 billion neurons. That's not the fact, the fact is that each of these neurons is actually equivalent to a five to eight layer, deep neural network. So did this modelling of like checking, if you take input output function of a single cortical pyramidal neuron, then how much how many layers and how many neurons do you need to be able to capture all of these input output mappings? And turns out, you need at least 1000 neurons connected in five years at least, to be able to capture it? So that's, that's like a window into the into the complexity of the brain. So now, my favourite part is it about the Pope.

1:33:54

I've been thinking

1:33:58

1:33:58

about the brain. It's also probably the most embarrassing things in neuroscience that we spend, have every day, a little chunk of our life, creating imaginary world, fantastic word in our head, when we sleep, and the fact that today was being saved in order to convince somebody you need the narrative, the fact that we are such narrative animal, and we are not just hallucinating with images we create amazing narrative will happen tonight, to all of us multiple time. And we know if nothing about it. We don't study it, because we feel embarrassed stuff, but after Freud to study it, so there is almost a stigma that you look a bit nasty if you studied. We don't even know how to study it, but we didn't even try to be perfectly honest. And this domain is fascinating fact and probably the most embarrassing things in modern neuroscience that such

1:35:00

Big Howard is called Institute of cognitive neuroscientists. This is a big cognitive things happen in whole our heads every single day. And we know so little about it. Good way to end. That's amazing. So just remains for me to say. Thank you so much for our panel. Thank you very much.

1:35:23

Thank you so much for our audience for coming and seeing us this went better than I expected anyway.

1:35:29

Thanks to all the producers and the people who are making this happen behind the scenes because there's a lot of them and it wouldn't work without them. So thank you