RODNEY BROOKS INTERVIEW PART THREE

Hello again, Ars Technica readers. This is the third and final installment of a three-part interview with Rodney Brooks – who is one of the most influential people in the worlds of robotics and AI. If you haven't yet heard part one or two, there are links to them on the page where this player's embedded, and I strongly suggest that you go back and listen to those installments before this one.

And with that - back to my conversation with Rodney Brooks.

Rodney Brooks: Yeah, let's talk about deep learning. Deep learning is the one that has enabled so much machine learning in general, it's on a lot of people's tongues. I saw an NFL ad recently saying we're going to bring machine learning to give you better diagnostics of players in real time with the Amazon Cloud. Deep learning is based on a technology called Neural Nets, which depending on how you look, the first paper was 1943, [inaudible 00:56:10]. It really became in the 60s, something that people were investigating. This famous book by Marvin Minsky and Seymour Papert called Perceptrons, which analyzes those neural nets. Then, in the late 70s to early 80s, there was a breakthrough algorithm called Back Propagation, which let the weights that are in these things, which are sort of vaguely related to synapses in the brain, let those be updated from the results and example.

Rodney Brooks: Back Propagation at the time completely overestimated in the short term. Thought it was the big thing. Back Propagation is one of the key aspects of deep learning and fairly soon people realized, okay, it can do certain things. It can't do as much as we thought. Most people moved on [inaudible 00:56:51] support vector machines, there are all sorts of other machine learning algorithms, flavor of the decade. A few people, Jeff Hinton at Toronto University, Yann LeCun who was there at that point, later moved to Bell Labs, later NYU kept pushing on Back Propagation. Then around 2012, it popped. Yann Lecun is chief scientist at Facebook now, runs an enormous lab worldwide for Facebook. Jeff Hinton is at Google and University of Toronto.

Rodney Brooks: It performs better than anyone expected. It was viewed as all those guys, they're working on that old problem. Good guys, they're making slow progress, but eh, that's never really going to go anywhere. It was one of a hundred things like that. We had no idea that Back Propagation was going to pop. We don't know whether any of those others are going to pop that way. In hindsight, there was no indication that it was going to be Back Propagation, which was the one that popped. There's no indication of how general it's going to be. The people who are seeing it for the first time, they think, "Wow! This is fantastic! [inaudible 00:57:55]"

Rob Reid: [inaudible 00:57:56] coming in 40 years. Yeah.

Rodney Brooks: I've people, people that you have interviewed, people that you have
                   interviewed say, "But, but. Don't things happen on a regular basis? Shouldn't we
                   expect the next one, the next one, the next one?" We don't know. We don't
                   know. We knew with Amara's Law that we were going to be able to continue to
                   have the feature size on a fairly regular basis for a long time. We knew that was
                   going to happen. People are trained to things getting exponentially better and
                   people think well, that's how everything works. No, science doesn't work that
                   way. Research doesn't work that way and all of these things take a long time,
                   like the self-driving cars. They don't just pop. This doesn't come out of nowhere.

Rob Reid: Then again, there's that timeframe distortion that you pointed to with self-
                   driving cars. Even as somebody who is fairly sophisticated in the field, my
                   perception is that deep learning has gone from zero to sixty in two or three
                   years, but it's over many, many decades. To quote something you wrote in one
                   of your blog posts, "Many people seem to think we'll continue to see AI
                   performance increase by equal multiples on a regular basis, but the deep
                   learning success was 30 years in the making, and it was an isolated event."

Rob Reid: Going back to your predictions, I think somewhat playfully, you predicted by
                   2020, the popular press will start having stories that the era of deep learning is
                   over. No earlier than 2021, VCs will figure out that foreign investment to pay off,
                   there needs to be something more than X plus deep learning equals profits.
                   Then, between 2023 and 2027, you predict the emergence of a generally agreed
                   upon next big thing in AI beyond deep learning.

Rodney Brooks: [inaudible 00:59:39] be many pretenders to the throne. You look at the
                   comments on my blog, lots of people said, "I know what it is, it's my research!"

Rob Reid: Of course. One of them will eventually be right. That's Amara's Law. Now, the
                   next one. Imagining Magic. Also hearkens back to one of your favorite quotable
                   sources, in this case, Arthur C. Clark who said, "Any sufficiently advanced
                   technology is indistinguishable from magic." How might that lead to a cognitive
                   distortion as people think about where AI is going and also about super
                   intelligence risk?

Rodney Brooks: We haven't seen any super intelligence. We have no prototypes. If you imagine
                   that it exists, if it's indistinguishable from magic, then it can do anything. The
                   example I like to use is suppose we had a time machine, why not? We transport
                   an elderly Issac Newton from his old times to now, but we do it inside the
                   chapel at Cambridge, he knew the chapel-

Rob Reid: [inaudible 01:00:33] building. He's like, oh, I'm here again! I know this place.
                   Nothing weird.

Rodney Brooks:	Yeah, make sure the lights are switched off and you have some candles around. Now, you pull out an apple and show it to him. The apple is an iPhone.
Rob Reid:	Yes. You don't drop it on his head this time.
Rodney Brooks:	No, you do not. But, remember, Newton, besides gravity, he figured out light. He figured out you could split up light optics and put it back together with prisms, so I'm really showing this iPhone. It's this new, flat device and then you press the side button, and the screen lights up with incredible detailed light. He has never seen a source of light that looks like that. Wow!
Rodney Brooks:	Now, you bring up an app and play a movie of an English Country scene, so it's just out in the fields with animals that he's seen, and it's playing on this thing. It's moving light. All he could do with his prisms was split it into colors. Then, maybe you go into iTunes and you play a piece of church music that he would have been familiar with. That machine can do that. Then you take photos of him and show him photos. You take a little movie and he sees himself. You turn on the flashlight and you show him in the dark corners underneath the pews that you can see stuff. It's a source of light and there are no flames.
Rodney Brooks:	Then, you go to the web and you find his personally annotated copy of his masterpiece, Principia, with his handwritten notes, and you can go on and look at every page. His own handwriting on his own copy is inside this little box. What limits does he understand about that box? What limits?
Rob Reid:	What could it not do?
Rodney Brooks:	What could it not do. He probably would be [inaudible 01:02:01] to realize that it's going to run out of power in a few hours. This powerful machine and it only works for a few hours, then you've got to do something else with it? That's pretty weird.
Rob Reid:	Yeah, 'cause he's never see anything that has that property before [inaudible 01:02:14].
Rodney Brooks:	It just is.
Rob Reid:	It just is. It doesn't just keel over and die.
Rodney Brooks:	So, he'll get that wrong totally. He would not understand that.
Rob Reid:	Now, he would certainly assume that this incredibly powerful could light a candle?
Rodney Brooks:	Yeah, it can do all this other stuff, and it's warm. It feels warm. He doesn't know what it can do or what it can't.

Rob Reid:	It seems omnipotent.
Rodney Brooks:	If you gave him a list of 10 things it could do, five which it could, and five which it couldn't, he would have no way of knowing which way the [inaudible 01:02:39]-
Rob Reid:	So, applying this to Super AI, we can imagine things it would almost surely be able to do, and extrapolating from that, we might imagine it could do anything.
Rodney Brooks:	Yeah. That's the argument that I get into with people when I talk about we're not going to have omnipotent AI anytime soon. They say, "You don't understand how powerful it's going to be," and these are people who don't necessarily work in AI. "You don't understand how power it's going to be." Neither do they, they have no clue. Just like Newton would have no clue of what it could do and what can't it, because it's so advanced that it's magic. It's nothing we've ever demonstrated anything close to, so he can't say anything rational about its properties.
Rob Reid:	Then, the boundaries vanish and we conceive of it as being functionally omniscient and omnipotent almost right out of the box? When in reality, if it were to start going down the path to omniscience and omnipotence, there would probably be many side journeys along the way.
Rodney Brooks:	Exactly.
Rob Reid:	The next one, what you call "Performance versus Competence."
Rodney Brooks:	I think this is perhaps the most important one. When we see a person perform some task, we have a generally good understanding of what their competence around that task is. If we see someone who plays chess better than anyone else in the world, we think they can probably teach people to play chess. We think they can probably explain to us why a certain move was important. What was their critical move in the whole game and why? Chess playing programs are better than any human and they can't do either of those things. The only way they teach people is-
Rob Reid:	By clobbering them.
Rodney Brooks:	By clobbering them, good game/bad game. You know? When we see some program labeling images, "Young people playing Frisbee in the park," if a person wrote that down in English, those very words, you gave them the image, they wrote that down. You would expect to be able to talk to them with questions like "What's the weather like in that picture? What sort of day is it? What sort of day is it outside right now? How far could a person throw a Frisbee?"
Rob Reid:	Or, what is a Frisbee?

Rodney Brooks:	Most important about a Frisbee, to be a Frisbee. You'd expect it to answer all those sorts of questions.
Rob Reid:	Can you eat a Frisbee?
Rodney Brooks:	Can you eat a Frisbee, yes. The person doing it could. They've got a competent understanding around Frisbee nerds if they can label something that programs don't know about weather beyond sometimes using weather words, because of some way the image appears, but they don't know what weather is. They don't know what a person is.
Rob Reid:	They don't know what a Frisbee is.
Rodney Brooks:	They don't know any of those things.
Rob Reid:	I'll certainly admit that when I see some of those eerily detailed mappings of descriptions of images that are coming out of the better image identifiers today, there's a presumption that when it says "A group of young people playing Frisbee in the park with a dog," that there is a wealth of understanding beneath that, but no. No notion of what a dog is, what a park is, what a Frisbee. The next one that I think is quite interesting is what you call "Exponentialism," the tendency to think that all tech, including AI, is exponential in nature, because we have been exposed to that so many times and so impressively in such a life changing way, with so many things in technology. Would you like to talk about exponentialism [inaudible 01:05:42]?
Rodney Brooks:	The example I use is from iPods, from the early part of the century where every 12 months, for the same price, they were roughly $400 then, they were coming out with double the memory, and they went from 10GB to 20GB to 40GB, and I projected at the time, and raised money for research of this projection, that by now we'd have some-
Rob Reid:	I think 160 terabytes.
Rodney Brooks:	Something like that.
Rob Reid:	Where we'd be by now.
Rodney Brooks:	[inaudible 01:06:06]
Rob Reid:	Of course.
Rodney Brooks:	The top of the line ones today are still 256GB. Why is that? Because, what was driving it was wanting to get people's music collection on them, and then eventually, the music collection, and there was no more driver.
Rob Reid:	There's market saturation.

Rodney Brooks:	Market saturation, so, it just stopped. We tend to think, oh, it's getting better, better, better. But, unless there's a market pool for it, it will stop.
Rob Reid:	Also, we often confuse S-Curves for exponential curves.
Rodney Brooks:	Right.
Rob Reid:	I mean, Amara's Law itself is probably an S-Curve, right?
Rodney Brooks:	Yeah, because we've gotten to the point where you can't have the size of the features anymore. You've got to do something different, which I think is actually a great thing [inaudible 01:06:38] computer [inaudible 01:06:38], but that's another whole story. Another example is GPS has gotten better, and better, and better, but at some point, you don't need sub-micron GPS to drive you care.
Rob Reid:	It's funny, because I made the same mistake that you did in looking at iPods, 'cause I was working in online music at the time. I remember wondering if I had done the right thing in creating Rhapsody as a streaming service, because it seemed unbelievably obvious to me that within 10 years, there'd be no need for streaming. You would just buy a device that came preloaded with all the music ever recorded, which would fit into a disc drive very easily, and you'd just get these wireless updates whenever Beyonce recorded a new song. I almost got to the point where I was like, "We've really blown it here. We're a little ahead of the curve with Rhapsody, but there's going to be no need for this." So, what are the ramifications for those who do think about the existential risk of a notional Super AI?
Rob Reid:	Very bright people like Stephen Hawking, Bill Gates, and Elon Musk, they have at least three things in common. One, they're all brilliant. Two, they're all very, very concerned about Super AI risk, and you pointed out that they have a third thing in common.
Rodney Brooks:	None of them has worked in AI. I respect all of them greatly, but I think they are making some of these other mistakes, that maybe Super AI [inaudible 01:07:52] in the future, but we don't know what the risks are going to look like. You go back to 1789, first hot air balloon is floating over Paris with people in it, and some people worried their souls would get sucked out, what's going to happen to those people up there? I don't think there was a single person in Paris on that day that worried about, we have to worry about noise abatement, where these things are ultimately going to land. That's going to be the limiting factor on these, how much noise they make. That was not the issue. That is an issue, that's why you get so much pushback against the extra runway a so many airports, but you couldn't tell that 220 years ago.
Rodney Brooks:	When I think these things are along the time scale, we can't begin to understand what the real issues will be. I think we should be much more worried about other issues that we have in tech. Why are infrastructures so leaky? What is it

that even our home thermostats can be used as attack vectors by putting viruses in them? I think there are much more immediate questions that come up. How is that fake news is going to effect our lives and our politics? I think there are a lot of issues that we see problems today. The problems about Super AI are way off in the future and we can't say anything sensible.

Rob Reid: There are a lot of very serious efforts underway right now to mitigate the risks that the future SUPER AI could pose to society. In Berkeley, there's an association called MIERI. The Machine Intelligence Research Institute. Elon Musk generously funded the Future of Life Institute. Would you say that they're wasting their time?

Rodney Brooks: I think they are. Some of these people you mentioned we have to regulate. Okay. Here's my question. If you're going to have a regulation and it doesn't change anyone's behavior, there's no point in having the regulation. There is only a point if it changes something. What is it you want to change? What is it? Tell me one example of what it is you want to change.

Rob Reid: I'm not sure how many are arguing for regulation. What I hear more about is what they call the Alignment Problem, the danger of a Super AI having and pursuing goals, which are inconsistent with humanity's wellbeing, and perhaps its future existence.

Rodney Brooks: Right, and the Alignment Problem has nothing to do with AI. The Alignment Problem is a real problem and the Alignment Problem exists in Facebook, it exists in Google, it exists in all these platforms. Those companies are founded by my friends. I'm not saying they're evil, but there's an alignment problem there, and it may, some may argue, destroy our democracy.

Rob Reid: I imagine some of the people who work in that field might respond something like this. Nobody claims to know precisely when this intelligence explosion will occur. Few of the most concerned people would guarantee with 100% conviction that it's even going to happen. Looking at recent history, and the compounding, and self-reinforcing effects of technological advancement, we can probably say that arbitrarily amazing things will be possible in the 50-100 year timeframe, maybe 200 year timeframe. Long for a human, but short for humanity. In light of that, what's the argument against working on this problem now? If now is not the right timeframe, what sign do we need? What development would it take for us to say, "Wow, now it's time to start thinking about this, working on this, or worrying about this"?

Rodney Brooks: Let me take two angles at that. One is I think we're more likely to see earlier than the pure AI, something which involves biological material. I think that's a much shorter way to get to renegade intelligence.

Rob Reid: Biological ... Oh, really? A development of intelligence?

Rodney Brooks: Because, you can build on intelligence that already exists. You reshape it some way, you're trying to build something, and-

Rob Reid: Supreme Machine Interfaces?

Rodney Brooks: Yeah, it could be around that or could be just a biologically edited animal-

Rob Reid: Or, organoids.

Rodney Brooks: Yeah. I think we're much more likely to see existential risks from them in the short term.

Rob Reid: You're not talking about bio weapons, you're talking about biological intelligence?

Rodney Brooks: Yes.

Rob Reid: Very interesting. I never heard that before.

Rodney Brooks: Yeah, I think that's much more likely in the short term.

Rob Reid: I'm personally very good at fretting about synthetic biology, but I never thought about synthetic biological intelligence.

Rodney Brooks: Animals are pretty damn intelligent. They're not that far from us [inaudible 01:11:55] on the evolutionary chain, so, take some existing animal, and you give it a few extra things. Who knows what that's going to be like. It could even be purely biological with the power of CRISPR and Design, and earlier [inaudible 01:12:10]. We don't really know what that looks like, by the way. You hadn't even thought of it.

Rob Reid: I hadn't, no.

Rodney Brooks: Then we get to an AI system that's got computers, and it's got sensors, and it's computational. One of the early warnings that we have to worry about, self-awareness, intentionality. An AI program for which tomorrow is different from today. There's no ongoing flow of time. For dogs, there's ongoing flow of time. Certainly for octopus', which are very different intelligence, evolved completely different from mammals. We don't have anything, anything remotely showing any of those signs. Until you can have dangerous AI, you have to have some sort of ongoing existence, some sort of ability to plan, some sort of ability to understand what's happening, some intention. We don't have any of those things, even in a rudimentary form. We don't have it at the level of an insect, so I'm not worried that we're close to it.

Rodney Brooks: Furthermore, just like you hadn't thought of the natural intelligence of an animal, I don't think we know what it's going to look like. Until we see some of

them, we won't know that sort's okay, or this sort is starting to look a little bad. Before we have robots that are really dangerous, we are going to have robots that are really annoying.

Rob Reid: Telemarketing bots!

Rodney Brooks: We're going to have all sorts of things along the way and I think we'll start to understand what the landscape looks like, and regulate as should be. We regulate all our other technologies, except for guns for some reason. So, it may be a fun game. It's like the Trolley Problem. Indistinguishable from magic. We don't know any properties, so we can imagine any properties we like, which is great for an academic wanting to write papers.

Rob Reid: If people want to work on the Alignment Problem, you would rather have them work on the alignments of things that are currently in our world and causing havoc?

Rodney Brooks: Causing havoc.

Rob Reid: Well, this has been an absolutely fascinating conversation. I probably could pose six more hours of questions to you, but I won't do that, because we've spent a lot of time together. Thank you very, very kindly for your time. Hopefully, at some point, we can reconnect and talk about some of the other amazing ideas that you've put forth in your blog, because there's many, many vectors we didn't even touch on.

Rodney Brooks: Yeah, this has been a really enjoyable conversation, so thank you.

Rob Reid: Thank you.

Rob Reid: Rodney's purview into tech's past, present, and future is remarkably deep and wide. As someone who has done a fair amount of fretting about the potential threats of Super AI, my recent 547 page novel being exhibit A, I take a fair amount of comfort from his sheer lack of concern about risks on this front, but not overwhelming comfort. I can't think of any other issue in high-tech, which divides quite so many brilliant minds, quite so vehemently. Rodney's of course correct that Super AI skeptics, including Bill Gates, Elon Musk, and the very recently, and very dearly departed Stephen Hawking, were never full-time denizens of this field. While that limits their direct expertise in it, it also leaves them un-conflicted in considering these issues.

Rob Reid: Upton Sinclair once said that "It's difficult to get a man to understand something, when his salary depends upon his not understanding it." Many of those who are most dismissive about Super AI risk are among those who could gain the most from its rapid and headlong development. That said, Rodney's own salary is in no way dependent upon advances in general artificial intelligence, as he's been a roboticist first and foremost for many, many years.

I'll repeat, the combination of his complete lack of alarm and extreme depth in this field gives me comfort, but I'll also repeat, not total comfort.

Rob Reid: I'm quite intrigued by Rodney's front line reports about factory worker shortages as far back as the 90s in what was then the poster child of cheap, abundant labor, China. I've done some digging since our interview and came across a recent-ish article in The Atlantic titled "China's Twilight Years," which says that the country's ratio of retirees to active workers will drop as low as 1.6 to 1 within about 20 years. That is an economy, which will need help keeping its factories humming. Particularly, if humanity doesn't get really good at creating robots that can help with eldercare, as Rodney so rightly pointed out, because factories are going to have to compete with that very important, and burning human need.

Rob Reid: Now, some listeners may dismiss a roboticist's claims about robots not imperiling factory jobs as being biased or self-serving. However, since our interview, I've tracked down a number of deployments of Rodney's robots on YouTube and elsewhere, and they generally seem to be in roles that enhance the productivity of the humans they share the factory floor with. This makes sense, because remember, these next generation robots are all about flexibility and re-programmability, and they need human hands and brains to pivot them from task to task, and to train, and tweak their actions. All of this reminds me of the concerns that a company on the rise of ATM machines, those little boxes seem to constitute an existential threat to the very title of "Bank teller," but over the years following their appearance, the number of tellers actually climbed significantly.

Rob Reid: The reasons seemed obvious in retrospect, which was with ATMs doing the simplest tasks, tellers started focusing on much higher value things, which made them much more valuable to their employers. It's a truism that when something becomes more valuable, we tend to buy more, or hire more of it. This was true for tellers after ATMs took over the mundane aspects of their jobs and it could be true for many factory workers who team up with robots like Rodney's.

Rob Reid: A third element of Rodney's thinking that really intrigues me is what he calls exponentialism. Living in these decades of steady, compounding improvement in computer performance has made all of us prone to this, but Amara's Law just doesn't apply to everything. Brief periods of rapid development followed by long periods of relative equilibrium are the rule in most dynamic fields. This has in fact been true about biological evolution for billions of years and more recently, we've seen this in air travel among hundreds of other industries. Planes got much faster quite suddenly when jet engines entered the scene, but they haven't sped up a whole lot since then.

Rob Reid: Punctuated equilibrium is common even in high-tech realms that we most associate with Silicon Valley. Consider Rodney's own field of robotics. From the time when the first factory robots appeared in a New Jersey car factory in the 60s until just a few years ago, substantially all industrial robots were caged, but

just a few swift years after Rodney's company rolled out its first product, thousands of uncaged robots are now strutting their stuff at major industry trade shows. Meanwhile, Rodney's point about the seemingly sudden emergence of neural networks and back propagation is an important lesson for all of us. A 30 year marathon can easily look like a three year sprint to outsides, significantly distorting

Rob Reid: I do hope that some of you will now join me on Patreon for quite a few extra thoughts about Rodney and his extraordinary thinking, because there's quite a lot to say. In this bonus content for Patrons, I focus mainly on several fascinating essays that Rodney has published on the internet. In preparing to interview him, I quite literally read every blog post that he has ever written, and I was awed by the wide range of topics that he tackles. In the Patreon recording, I talk about the pieces that fascinate me the most. The episode runs almost 20 minutes, so, I guess you can kind of think of it as the last quarter of this week's podcast. Once again, to access that, go to Patreon.com/RobReid R-E-I-D. If you support the podcast at the $5 per month level or above, you can hear that, as well as all the extra segments that I posted for other episodes.

Rob Reid: Thanks for listening and I hope you'll join me and George Church next time.

END INTERVIEW ELEMENT OF PART THREE
So Ars Technica listeners - here we conclude the third and final installment of my interview with Rodney Brooks. Thanks for sitting through that brief description of my Patreon feed at the end.

That was recorded back in March. And if you're interested, a couple things have since changed about my Patreon feed. And if you're not in please join me back here on Ars next week and feel free to turn this off now. But if you're interested, I enabled this really cool feature on Patreon which let me set up a private podcast feed for the people who subscribe to my extra bits of content. So my Patreon extras flow right into your smart phone, as if it's a separate podcast that you're subscribing to. So to hear something, you don't have to go to the Patreon site, log in, and hit play. So now you can hear this extra stuff when you're on the go, just like any other podcast. It's also a separate feed, so you don't get confused about what's a main episode, what's extra. It really works great, and hats to Patreon for building that.

The other slight update about Patreon since I first posted this interview is that I've now done bonus segments for ten episodes, so there's now hours and hours of stuff up there. And it's always a little bit different. Sometimes I interview a second expert about the interview that's in the main podcast feed. Sometimes I pull together a very high-density set of follow-on thoughts & conclusions - which is what I did for the Rodney Brookes episode. And there's actually a ton of really cool Rodney-related stuff that we didn't have time for in the episode itself

And this week I did something complete different & fun, for my interview with Stewart Brand, which is currently live in my main podcast feed as I record this. For that, I posted a highly extended version of the interview for my patrons. So while the interview in the public podcast

feed is about 90 minutes long - which isn't short to begin with - the Patreon version is well over two hours. So it's about 50% longer, and it's really lively.

I'm guessing many of you know who Stewart Brand is. If not - if you heard last week's episode here on Ars with George Church - Stewart is the guy who's working with George to resurrect extinct species like the woolly mammoth. He has also lived a wildly impactful and influential life in tech, in science, and in culture going clear back to the early 1960s. So it was one of the most amazing conversations I've ever had, with someone who's frankly with one of my heroes. You can find it at after-on.com right now.

Anyway, I hope you'll join back here on Ars next week. I'm bringing another episode from my archives. It's one of my favorites, it's a conversation with Mary Lou Jepson, who is an amazing engineer, entrepreneur and thinker. If you're interested in neuroscience, in you're interested in holography, if you're interested in telepathy -- and who is NOT interested in one off those things. I believe you'll enjoy it a ton.

I hope you join us.