Humanity Enhanced: A Conversation with Ray Kurzweil

Speaker:
Ray Kurzweil, author of The Singularity is Near, and Director of Engineering, Google

Interviewr:
David Kirkpatrick, Techonomy

Kirkpatrick: Well, next, we're going to have somebody who I'm very excited about, Ray Kurzweil. Somebody who I'm sure many of you are familiar with. I'm going to start introducing him now, even before he gets on stage.

He was called the ultimate thinking machine by *Forbes Magazine*, which is one of our sponsors so we like the Forbes quotes. And then *Inc. Magazine* ranked him number eight in the history of the U.S., calling him the rightful heir to Thomas Edison.

I'll just read you a list of some things that he has invented or been a principal developer of. The first CCD flatbed scanner, the first Omni-font optical character recognition, the first print-to-speech reading machine for the blind, the first text-to-speech synthesizer, the first music synthesizer capable of recreating the grand piano and other orchestral instruments, the first commercially marketed large vocabulary speech recognition.

Ray, I hope you're out there and can join me on stage. But I could keep going. Winner of the $500,000 MIT Lemelson Prize, the world's largest for innovation. National Medal of Technology from President Clinton.

Anyway, please sit down, Ray.

Kurzweil: I didn't want to interrupt.

Kirkpatrick: And an author, on top of all that, who's written a couple of really well-known books, including *The Age of Spiritual Machines*, *The Singularity Is Near*, and now, a new book, which isn't even out yet, called *How To Create a Mind: The Secret of Human Thought Revealed*, which you have all received today, as you have registered, a week before it's publicly available. So we're very pleased about that. And Ray made that possible. So we thank you for that.

So welcome.

Kurzweil: Great to be here.

Kirkpatrick: Let me see if I can find the questions I wrote down for you. Okay. I don't know if you were listening carefully to the previous panel, but, you know, we were talking about the ways the world is going to be changed. And I wondered if you thought that we went far enough or, you know, you have some pretty extreme views, in the opinion of many, about how dramatically human life is going to be altered.

Now, that's a longer-term view. But how optimistic are you sort of as you look at all the trends that you have studied so carefully and participated in making possible, how optimistic are you about sort of the next 5 or 10 years on the globe and our ability to start solving problems at scale?

Kurzweil: Very optimistic because there is a lot of evidence now that not only hardware is progressing exponentially but software. Some of my critics say, well, Kurzweil's right that hardware is progressing exponentially, which by the way, in the case of computers, goes back to the 1890 Census, but software is stuck in the mud.

And there's many ways of measuring this. I talk about that in *Singularity is Near*, and in this new book. Recently, the scientific advisory board to Obama looked at this question of hardware versus software contributions. And they looked at every major

scientific and engineering problem area. Like, for example, linear programming, that's a certain type of engineering technique, which I actually believe is used in the human brain itself. And they came up with the conclusion that there's been improvement over the last 15 years of 45 million to 1; 1,000 to 1 from hardware and 45,000 from software. They multiply together to get 45 million. In each case, they find more contribution from software than hardware.

And we see viscerally impressive examples like Watson, which people are quick to dismiss, but what they don't understand is --

Kirkpatrick: The IBM computer.

Kurzweil: Yeah, the knowledge that Watson had that that 16th Century Dutch princess with blonde braids did not come from that being coded in some list-like language. But Watson actually got that by sitting down, so to speak, and reading Wikipedia.

Kirkpatrick: This is one of the winning questions on Jeopardy that allowed it to win? Is that what you are referring to?

Kurzweil: It had some very impressive questions. In fact, I brought one. It got its information from reading Wikipedia and several encyclopedias. 200 million pages, like, for example, a long, tiresome speech delivered by a frothy pie topping.

Kirkpatrick: I'm not going to answer.

Kurzweil: What is a meringue harangue, which Watson got correct.

Kirkpatrick: A meringue harangue, it figured that out. That's appropriate, good.

Kurzweil: It got quite a few like that. It got some easy ones wrong. But it got a higher score than the best humans put together.

Kirkpatrick: So going back to the point about software then, is that, again, to the 5- or 10-year question, if software is increasing that quickly, it implies our ability to wrestle with big, heretofore unresolvable problems in human health and things like that?

Kurzweil: There is all this fantastic knowledge in millions of books and Web sites which heretofore we have been able to kind of categorize by key words. Amazing how useful that is, but we'll actually be able to have computers go out and read that and understand it. Watson now is reading all medical literature, for example, to provide a very comprehensive medical diagnostician. There is knowledge to be gained in all of the written language that we have assembled.

Just one example. The Google self-driving cars. That would have seemed like science fiction 10 years ago, and yet that's a reality. I have talked to people who have ridden in them and they quickly gain more confidence than in human drivers, which is maybe not saying much.

Kirkpatrick: Now, one of the things you said when we talked on the phone was in coming years we'll be doing a lot of our thinking in the cloud, which I thought was an interesting formulation. Is that another way of sort of saying that we will all have access, so to speak, to Watson in whatever we do?

Kurzweil: Well, that's kind of step one. I mean, right now if you do anything interesting with our devices, if I do a search or a language translation, I'm going out to the cloud. It doesn't take place in here. In fact, during that one-day SOPA strike, I felt like a part of my brain had gone on strike.

But we literally would be expanding the essence of our thinking into the cloud. If you go out to the 2030s, what I talk about in this book is how the neocortex works. In summary, it's 300 million panel recognizers that are all pretty much the same, except that what they are able to do is to wire themselves into these elaborate hierarchies.

So at a very low level, I've got little panel recognizers that go: Oh, crossbar and a capital A. At a higher level, one will go off: Oh, the word "Apple." At a much higher level, one will go off saying: She's pretty. That's funny. That was ironic. You probably think those are much more complex. They are actually the same. They just sit at a higher level of this hierarchy.

Where does the hierarchy come from? That is actually created by the neocortex from our own thinking. So we are very much what we think. So be careful who you hang out with.

But we are limited to 300 million. That was enough on the one hand --

Kirkpatrick: 300 million what?

Kurzweil: Pattern recognizers.

Kirkpatrick: Pattern recognizers.

Kurzweil: Little modules.

Kirkpatrick: In the brain.

Kurzweil: Neural Net has been the wrong approach because it has one neuron as the basic element. One neuron can't really do that much. The basic element is about 100 neurons and it has a fair amount of sophistication and it can recognize a pattern in complex ways, even if parts of the pattern aren't there. And it can wire itself in these hierarchies. Only the neocortex can do hierarchal thinking, and only mammals have a neocortex.

300 million, is that a big number or a little number? Well, the big innovation in Homo sapiens is we have this big forehead to allow more neocortex. And that quantitative advantage was the enabling factor for us to create language and music and art and science and technology. And no other species did that.

But it's also limited. And if you struggle trying to learn a new skill or new language or even just read a new book, we run up against the limitations of those 300 million pattern recognizers. We fill them up by age 20. Can't really learn anything new without forgetting something old. That's actually an important skill to have.

But ultimately, we'll be able to expand it. We'll send nanobots. I mean, these are getting smaller, at a rate of 100 in 3D volume per decade, and these will be the size of blood cells in 2030. So go inside our brain and basically expand our neocortex into the cloud because that's where interesting things happen.

Kirkpatrick: But in the short-term, can't we move in that direction with the clouds held on our apps and access to massive data sets and --

Kurzweil: Absolutely. This is basically a window to the cloud now. Anything I do that's interesting, it takes place in the cloud. And even within a few years, you'll have search engines in the equivalent, other knowledge-finding tools that have actually read the 10 billion pages on the Web and have read millions of books and understand, to some extent, not at the human levels but much more than they do now, which is basically not at all, the content. And just the way Watson is able to understand pretty complicated content.

Kirkpatrick: So will doctors potentially be rendered in many cases unnecessary once Watson or computers like it have acquired this information?

Kurzweil: Watson is going to actually—that kind of system is going to become a very reliable tool which people will get dependent on. Just the way the people I talk to have ridden in the Google self-driving cars and people ask them: Well, weren't you nervous? They actually quickly get more confidence in the AI driver than in human drivers. And these AI systems will have

read every single journal article and medical blog of significance, which no human doctor can do. So they will actually be a master repository of our accumulated human knowledge.

There are still tasks for humans to do as long as there are certain ways in which human intelligence is superior, which it is in terms of in particular relating to other humans and understanding human emotion, being funny, getting the joke, being sexy, expressing a loving sentiment, you know, being sympathetic. These are not sideshows to human intelligence. That's the cutting edge of human intelligence.

And humans are still superior. My view, though, is we are creating computers in our own image to basically make ourselves more of what we value in human intelligence. And we're going to be merging with that technology. We already have.

Kirkpatrick: But given what you just said, you also have said, you know, that computers of the future will be able to write convincing poetry and novels. I mean, is that contradictory to what you just said about the distinctive uniqueness of the human being? Or is it just a way of saying there will be poetry and novels, but they probably won't be as sensitive and passionate and heartfelt or as powerful as a human-written version of the same thing?

Kurzweil: It's not contradictory because I consider those AIs of the future to be human once they can write a convincing novel.

Kirkpatrick: So you really would call them human at that point?

Kurzweil: Yeah. I mean, that is the essence of a Turing Test. And I've been consistent in saying the computers will pass the Turing Test by 2029. Meaning they really will be convincing in their ability to do human-like things. And if you have a computer that passes the Turing Test but isn't convincing in that way, then it's just a badly designed test, because Turing wasn't very precise on how to set up that test.

But it is exemplary of our civilization. We are human machine civilization. And computers are doing things all the time we couldn't possibly do. Mathematica greatly excels at math compared to human mathematicians. And this AI does not exist in a few dark government intelligence laboratories. It's very widely distributed. A kid in Africa with a smartphone has access to more intelligently searchable information than the President of the United States did 15 years ago.

Kirkpatrick: I love the kid in Africa continuing to come up. It is a very good metaphor, or reality, actually.

Kurzweil: It's happening, yeah.

Kirkpatrick: Will the audience for those poems and novels be other computers that are themselves effectively humans? Or will they be us? Will we be moved by the poetry written by a computer that you call a person?

Kurzweil: If we're not, then it's not operating at human levels. I mean, ultimately AI will operate at the best human levels. So when you say, “is a computer doing music as well as a human?” what does that mean? Is that the guy whistling in my hotel room this morning? Or is that Beethoven or John Lennon? There's a very broad range of human ability. Ultimately computers will operate at the best of those levels. But we're going to use those tools to make ourselves more expressive and more intelligent.

Kirkpatrick: Okay. Talk about the book and let's hold it up, just to be only fair, since you have come this long way and given it to all of us. *How to Create a Mind: The Secret of Human Thought Revealed*. What makes this different from your other books? And why did you write this book now?

Kurzweil: Well, I've been thinking about thinking for 50 years. In fact, I wrote a paper 50 years ago, when I was 14, that the essence of human intelligence was pattern recognition.

Kirkpatrick: Like the rest of it. 14, paper, whatever. Yeah. Go ahead, sorry.

Kurzweil: So this was a high school science paper I submitted to the Westinghouse science talent search and got to meet President Johnson. And I have the same view, except I can articulate it more clearly.

The best evidence has just come out. In fact, I was about to send this off to the publisher. And four times there was some great study that really supported my thesis that I held up, no, I've got to include this. And the publisher was getting very frustrated. We actually pushed back the date because of that phenomena.

But it's only recently that we can see inside the brain with enough precision to see individual interneural connections, see them forming and firing in real time. We can see our brain create our thoughts. We can see our thoughts create our brain. And drawing from neural science, from actually some simple thought experiments, which I guide the reader through, and some of my own work and the work of others in AI which only shows which techniques could possibly work; and drawing upon all of those sources, I describe a thesis for how the neocortex works, which is that it has a uniform algorithm. And it really runs contrary to a lot of neural science, which is very fond of saying: Oh, well, this region does this. It's the fusiform gyrus [that] recognizes faces and V1 recognizes the edges in objects, because we see [that] if those regions get knocked out, suddenly people lose those skills. Although they can relearn them.

One of these pieces of evidence that came out just as I was sending off the book is what happens to this region V1 which is considered to be a very low-level, primitive pattern recognizer that recognizes edges and shadings, basic features of visual objects in a congenitally blind person who is not getting any information? Does it sit there and do nothing?

Kirkpatrick: So some empirical research was done on that?

Kurzweil: Yes. It actually gets harnessed by the frontal cortex to do high-level concept recognition in language, which is at the opposite of the spectrum in terms of level of abstraction of concept.

So even these very highest level concepts, like that's ironic, turn out to be the same basic recognizer. And the neocortex is very interchangeable in terms of its application. When air gets knocked out, another area --

Kirkpatrick: So one of the key insights is the uniformity of functionality across that part of the brain.

Kurzweil: Right. And then from other experiments we can infer what that functionality must be. And a key is that we can actually recognize and think in hierarchies. Only the neocortex can do that and only humans can do it at the level we can do it. Before mammals, there was no neocortex and animals could only learn over thousands of years through thousands of generations of gradual biological evolution to adapt to new circumstances.

So 65 million years ago when the crustaceous extinction event happened, the sudden—you know, maybe over several days, cataclysmic environmental effect, the non-mammalian species without a neocortex couldn't adjust. That's when the mammals took over.

Kirkpatrick: Let me just make sure I understand it. The significance of wanting to write a book about this theory and set of discoveries about the neocortex is what exactly? Is it that it implies the ability to essentially extend intelligence by attaching other computational processes that are external because of what you have discovered about the way the brain itself works?

Kurzweil: It's a grand and somewhat diffuse project to understand the human brain which people are not fully aware of and is perhaps the most important project in the world. And there are three reasons for it, or purposes. One is to be able to fix the brain better. So, you know, early examples like the neural implant for Parkinson's disease does not treat the brain as a chemical soup like SSRIs do. It actually goes right in and tries to fix certain connections. And the latest version of that actually allows you to download new software from outside the patient. The *New York Times* had an article recently expressing concern about people hacking into the software that people are downloading into their brains. But anyway, fixing the brain is one based on a better insight as to how it works and better tools to intervene is one purpose. Another purpose is to provide the models to

create more intelligent machines, AI. We have not actually benefited that much from insight from neuroscience in AI because we haven't really been able to look inside the brain with enough precision. However, ironically, the technique that has taken over many parts of AI—it's a technical technique called hierarchical hidden Markov models, something I helped pioneer in the '80s and '90s that took over speech recognition and much of natural language and understanding. And it's used in Watson, happens to be similar to the mathematical technique that the brain uses. They both evolved kind of independently because they work, not because in AI we understood that that's how the brain works.

But as we get really better models for how the brain does intelligent things, we'll do a better job of creating artificial intelligence to augment our own intelligence. And finally, it gives us greater insight into ourselves. I mean that's the whole purpose of both art and science—is to understand ourselves.

Kirkpatrick: Okay. Well, that's good. I want to take audience questions and comments. And we can take tweets if any come in that are pertinent. Anybody have a question or a comment for Ray? Over here. Oh, Stephie, please. Good to see you here. Get the mic real fast before—Could we get the mic real fast to Stephie over here? Stephie came from Munich to Techonomy. Thank you, Stephie.

Stephie: Ray Kurzweil, hi. I'm very interested, Mr. Kurzweil. Are you an optimist? Or are you a pessimist? Are machines taking us over?

Kurzweil: I have been accused of being an optimist.

Kirkpatrick: I think I accused him of that earlier actually.

Kurzweil: There are two issues. There's optimism on feasibility and desirability. They are really two completely different issues. I am an optimist on feasibility, although there are—there's a camp of people that actually think I'm too conservative at 2029. The consensus in the AI field.

Kirkpatrick: Referring to passing the Turing Test in 2029.

Kurzweil: Yeah. The consensus in the field has been moving towards me. I was at 2029 15 years ago, and the consensus in the field was hundreds of years. And at the AI 50 conference on the 50th anniversary of the 1956 Dartmouth Conference that gave artificial intelligence its name, the consensus was 50 years. And I was closer to 25 at that time. Now it's people are saying, well, maybe 2030, 2040, 2050. I mean, we're growing closer, but not because I'm changing my position. But I'm considered an optimist on feasibility.

On desirability, I'm considered an optimist, too. However, I have written probably more than most other observers on the down sides. Technology is a double-edged sword. Fire kept us warm, cooked our food, but burned down our villages. Biotech today, we're working on reprogramming biology away from disease. It's a whole other story that's also progressing exponentially. But the same tools could be used by bioterrorists to create a bioweapon. I've actually worked with the Army on addressing that.

AI, I mean we've all seen dystopian movies where the theme is that the AI is unfriendly, to use a technical term from the field.

Kirkpatrick: And that's not crazy.

Kurzweil: And that's not crazy at all. In fact, if you have an entity that's more intelligent than you are, bent for your destruction, that's not a good situation to get in. The only way out of that would be to get another AI that's even smarter than it on your side.

Kirkpatrick: Sounds like you could write a science fiction movie, yeah.

Kurzweil: Well, I'm actually working with Roland Emmerich on a movie called *Singularity*. But the answer I think—I take some comfort from the fact that the AI today is very widely distributed. It doesn't exist in a few unstable laboratories. And it is empowering all of humanity and is part of human civilization, which is a human machine civilization. And the best way to reflect the values that we cherish in the future is to practice them today. And people say, well, that doesn't sound like a very foolproof strategy, and it's not. But that's the nature of history.

Kirkpatrick: Okay. Over here. Mark. I don't know the names of everybody who's going to ask a question.

Bonchek: Thanks, David.

Kirkpatrick: Identify yourself, though.

Bonchek: Mark Bonchek, the chief catalyst at ORBIT+Co. When you were talking about the brain as this network of hierarchies of 100 neurons each, I was struck with the similarity of Dunbar's number of 150 and the way our organizations work. And I'm curious if you see a comparison between the brain and organizations as essentially networks of hierarchies, of pattern recognizers, aggregating up to have some kind of collective purpose. And will organizations change as we understand how the brain works?

Kirkpatrick: That's a good question, Mark. Nice. Okay. Let's hear the answer.

Kurzweil: There is a metaphor of that there. Minsky articulated in the name of his book *Society of Mind*. And some of these societies of mind are better and more effective than others, the same as we see in human societies. And so a totalitarian society is not going to be as good at making decisions, let's say, as a well-balanced, well-communicating Democratic society. And we see—even though we all have 300 million and we all have the capability of wiring it up in a similar manner, it's all based on every little decision we have made and everything we are exposed to. And obviously some people have a mind that's more effective than others, at least for certain tasks. I mean Beethoven was great at music and Einstein was great at metaphors and physics. So that's something else we can study. I think there are metaphors there. We extend actually the 300 million panel recognizers by forming a society, and a group of people through collective decision making can solve problems and individuals can't.

Recently, several hundred mathematicians got together, never met each other, using these collaborative online tools to basically form a supermind and were able to very quickly dispense with, that is to say, solve, a mathematical problem that had been unsolved for several hundred years. So we do have ways of actually making ourselves smarter through collaboration. That was the value of generating verbal language and written language, our first two inventions. We have ways of combining minds that don't amplify our intelligence, like the phenomenon of a lynch mob. So this is a very important area to study. I think there's a metaphor there, but the mechanisms are somewhat different.

Kirkpatrick: Unfortunately, we have to wrap. This has been extremely interesting, Ray. I'm so pleased that you could join us.

Kurzweil: Good to be here.

Kirkpatrick: Thank you again for giving us all copies of your book, for coming to Tucson. Keep up the good work and let's get there sooner rather than later. Okay?

Kurzweil: Thanks so much.

[APPLAUSE]