

# **SDS PODCAST EPISODE 241: PUSHING THE BOUNDARIES IN MENTAL HEALTHCARE WITH DATA SCIENCE**



- Kirill Eremenko: This is Episode Number 241 with Dr. Guillermo Cecchi from IBM Research.
- Kirill Eremenko: Welcome to the SuperDataScience Podcast.
- Kirill Eremenko: My name is Kirill Eremenko, Data Science Coach, and Lifestyle Entrepreneur. And, each week we bring inspiring people and ideas to help you build your successful career in data science. Thanks for being here today, and now let's make the complex, simple.
- Kirill Eremenko: Welcome back to the SuperDataScience Podcast. Super excited to have you on the show and today we got a very special guest, Dr. Guillermo Cecchi from IBM Research. So, what did we talk about in this podcast?
- Kirill Eremenko: Well, Dr. Cecchi specializes in computational linguistics, and over the past couple of years has pioneered approaches to quantify psychiatric and mental conditions from speech samples. A very exciting field and Dr. Cecchi's research is applied to conditions as diverse as schizophrenia, mania, prodromal psychosis, drug and alcohol intake.
- Kirill Eremenko: In this podcast, you'll find out how data science and artificial intelligence are pushing the boundaries of healthcare and specifically mental healthcare, and what can be done there. Some very interesting approaches which extract data and insights from audio samples of patients' voices and their speech.
- Kirill Eremenko: You'll also learn about a couple of interesting techniques. Specifically one about transferring

intuitive knowledge that professionals in this field have into algorithms.

- Kirill Eremenko: And, finally, towards the end of the podcast, we'll have a bit of a debate about technology and the brain, and things like Neuralink, and evolution, and where the world is going in the next couple of years.
- Kirill Eremenko: So, quite an exciting podcast, especially if you're in the space of science, or you're interested in things like how the brain works, and natural language processing and it's application in healthcare.
- Kirill Eremenko: So, on that note without further ado, I bring to you, Dr. Guillermo Cecchi from IBM Research.
- Kirill Eremenko: Welcome back to the SuperDataScience Podcast, ladies and gentlemen. Super excited to have you on the show. Today I've got a very special guest with me calling in from New York, Guillermo Cecchi.
- Kirill Eremenko: Guillermo, welcome to how are you doing today?
- Guillermo Cecchi: Very good, warmer than last week.
- Kirill Eremenko: Yeah, we chat about that just now. How cold did it get last week?
- Guillermo Cecchi: Minus 16, minus 18 Celsius, so. Yeah.
- Kirill Eremenko: What is that in Fahrenheit?
- Guillermo Cecchi: Oh, one day was close to one or two Fahrenheit. Three, five, I think it was, yeah. It was something.
- Kirill Eremenko: Did all of New York stop for that time?
- Guillermo Cecchi: No.

- Kirill Eremenko: No. Yeah, I was talking to a friend of mine in the UK and they had like a similar situation where they actually had snow. I think it was last week. And the whole city just came to a halt, because they are just not used to having snow there at all. Yeah. Crazy weather these days.
- Kirill Eremenko: Alright, well Guillermo very excited to have you on the podcast. You're with IBM Research and you're in an extremely exciting area which is research of the brain, and psychology, psychiatry. And in the connection with technology, data science, artificial intelligence. So super, super pumped about that.
- Kirill Eremenko: Could you give us a quick rundown of your background? Because you started all the way in physics, and now you are in data science, artificial intelligence, and brain and psychiatric research. How did that go? How did your journey unfold? Could you please give us a quick overview?
- Guillermo Cecchi: Well, yeah, it's really there were no right ... I started as a physicist, but I always had a keen interest in philosophy. And I actually before coming to the US to do my Ph.D. in physics, I was taking master courses in philosophy. So that naturally led me to brain science. And in my Ph.D. while I was doing things in physics I was actually working at the interface between physics and biology, and then physics and neuroscience. And then I did a postdoc in psychiatry at Cornell Medical Center, also in the city.
- Guillermo Cecchi: And when I joined IBM, I joined a group that was doing brain inspire [inaudible 00:05:38] models so I kept

working also in neuroscience. And then at some point, I had the opportunity of getting access to language data in particular, from mental health patients. And this is when I started to dedicate most of my time to this.

Guillermo Cecchi: Right, so the background, it's what led me to this conference of psychiatry, neuroscience, and data science, mathematics, computer science. So being an IBM, naturally, I'm surrounded by computer scientists so that's a given. So that's, in a nutshell, the background.

Kirill Eremenko: What I find really interesting is the way you moved from physics through some exploration of philosophy. You moved into a postdoctoral in psychiatry. So that from my perspective that requires a lot of courage that completely changing your direction in life.

Kirill Eremenko: Just could you comment on that? How did you feel about that move? Were you nervous that you're completely shifting your career? Or were you excited and not even thinking twice about getting into something that you can see that's your passion?

Guillermo Cecchi: Well, I was excited because I realized that there is so much to do. I was never in doubt that I will find my path. Just knowing how much we don't know about the brain, right? So [inaudible 00:07:46] can be daunting, at the same time, it can be exciting because everything has to be done. So, in that sense, I was not scared at all. I was very excited.

Guillermo Cecchi: There is so much uncertainty. When being a postdoc it's always difficult, because it's like being an

adolescent in science. When you're a child, you're protected. When you're a faculty, you have some protection. But being a postdoc is probably the most difficult position in science.

Guillermo Cecchi: So, but even then, I was not anxious scientifically. Because I knew as I know that there is so much to do.

Kirill Eremenko: That's some great advice because like we have lots of listeners who are interested or curious about the field of data science but they're actually in a different field right now themselves. Where they might be a developer or they might be more into the statistics side of things or they actually might be in a completely unrelated field, a creative field. And this is a great testament that indeed there is so much to do.

Kirill Eremenko: When there is so much to do and when you're really truly passionate about the field ... I like how you said that there is, you never had doubt that you will find your path. Sometimes a path isn't clear you can't see it, but you just don't have the doubt that you'll find it. So I think that's some inspiration for our listeners.

Guillermo Cecchi: But, yeah, this, that, whatever your background is in this field of neuroscience, psychology, psychiatry. There is so much to do, and there are so many things that you will find something that will be your passion. Because it touches upon everything.

Guillermo Cecchi: Every human concern passes through the brain. So, it doesn't matter what tools you have with you. You will find the connection with this field because it is important for all of us. So if you talk about mental health, who doesn't have one person in their family, or

a friend who suffers mental health? Like that, you can find a personal connection with this field for everyone.

Kirill Eremenko: Yeah, that's absolutely true indeed, mental health is ... you know like, I think, indeed, just as you said, maybe a family member or friend we all have somebody we know who might be going through some issues. But even on an individual level, I find that we go through cycles as humans. We feel good ones-

Guillermo Cecchi: Absolutely.

Kirill Eremenko: -and bad like, Why? Why does that happen? Why does our brain ... I was watching one of your videos and I loved one of the quotes that you said. I think this is absolutely genius. You called the brain these two pounds of protein and water.

Kirill Eremenko: Why do these two pounds of protein and water do make us feel like that, or like this? Why do we have mood swings? Why do we have good days and bad days? Why do we feel attracted to certain people and not attracted to others? Or passionate about certain things and not passionate about others? Those things are still all unexplored. It'd be really interesting to understand how all that works.

Guillermo Cecchi: Yeah.

Kirill Eremenko: Yeah okay well that's very exciting. Now what I think would be important to mention, at this point is, how is data science in these days and technology in general, how is it helping in medicine? What is the link? What is the connection? So, for the benefit of our listeners, for instance, people who are into data science but have



never worked in the industry of healthcare or medicine, what is the potential for them to apply their skills in this industry?

Guillermo Cecchi: Right. Well, of course, there are many directions. Just thinking of traditional machine learning or AI approaches. There is such a need to combine information that is present in medical records, for instance, hospitals have this extensive medical records going back decades that for the most part are just sitting in servers. If not, in drawers.

Guillermo Cecchi: And we have 46 people here in the lab working on that. Just going over those records with current computer science techniques. And extracting patterns that are very relevant for understanding the progression of the disease. Or understanding the probability that someone who came to the hospital a few times will have a certain event in the near future.

Guillermo Cecchi: Respect to, I can give an example regarding what we do in my group. So for instance, we have the paper coming out and we have presented this in conferences where we have patients who have Parkinson's disease. They take a particular drug Levodopa, it's something that improves their symptoms in many cases. Not everyone responds to the drug but many people do. And you want to know what the effect is and what is the best time to take the drug for them.

Guillermo Cecchi: And you as a neurologist, you can see the patient once a month. In extreme cases maybe once a week. What happens in between with a patient? How do you reach the patient on a daily basis? So, we showed in this



paper that just by recording their voice we can detect whether they have the effect of Levodopa on, or they're off Levodopa. And with that we can track effect throughout the day by just recording their voice in a cell phone. It's very simple technologically, and the people who actually worked on this and obtained the results at the moment had no background whatsoever in neuroscience, psychiatry or neurology. They were computer scientists, electrical engineers and just by applying data science to extract features from the voice, and learning along what it was known in the literature that but it was not formalized.

- Guillermo Cecchi: They were able to come up with this result and this is direct application that has immediate impact. Something that we can use to turn around the traditional ways of doing mental health care by empowering these tools through a modern Information Technology. Really being able to access, to have a picture of your patient on a daily basis. That's something that otherwise will be impossible. So I think that that's an interesting good example of how bringing your knowledge in data science and analytics, you can have an immediate impact in this field.
- Kirill Eremenko: Okay, so just from the voice of a patient you can extract certain insights about his mental health? Is that right?
- Guillermo Cecchi: Yeah so in this case, you can think of in a simple way. If you're taking a psychoactive drug, so a drug that has an effect in your central nervous system naturally, your voice would be affected. So the nervous system

controls your muscles and, of course, the muscles that control voice would be affected.

Guillermo Cecchi: But also, we extract information from language from the lexical content. So we study information not only from the voice, and frequencies and the prosody that from the content so that, again, your central nervous system is affected. So of course, you might expect this. But the fact that these Parkinson patients are on Levodopa, so their symptoms are improved.

Guillermo Cecchi: It changes, for instance, the way they retrieve certain words. It's more likely that, for instance, they will talk about verbs related to action and movement than when their symptoms, which are manifested among other things in motor control, when their symptoms are high they're more Parkinson that tendency disappears.

Guillermo Cecchi: So just using in data science techniques we can capture that and measure it and provide a metric of whether the drug is having the positive effect or not.

Kirill Eremenko: Okay. Gotcha. Gotcha. So, and can you talk a little bit about IBM Watson here? Because I think that's another great example of how technology can be used in healthcare. So a couple of years ago we heard some breakthroughs where IBM Watson beat humans at the game of Jeopardy.

Kirill Eremenko: Then there were some example case studies of how IBM Watson can predict certain things and actually better than doctors and faster than doctors. Because it has access to much more literature, much more

materials, and case studies. What's going on with IBM Watson these days?

Guillermo Cecchi: Well, I cannot tell you much because that is let's say the commercial offering part and we are research. So what goes into Watson, in Watson Health in particular, are tools that have been solidly tested and questionable from the point of view of the clinical relevance. So what we are doing is, we're creating the pipeline of the things that eventually will end up in Watson, in Watson Health in particular.

Guillermo Cecchi: We use some tools, the speech recognition, for instance, we use tools from Watson. And some of the machine learning platform, built in Watson we also use. But what we are doing here in research is it's exactly that, we are doing basic research. And out of that, what is really stable will eventually go into Watson.

Kirill Eremenko: Okay. Gotcha. So, you're kind of like pushing the boundaries and exploring new ways.

Guillermo Cecchi: Exactly.

Kirill Eremenko: Okay, okay. Very interesting. Alright, so we talked about two main applications. So extracting data from medical records that there's plenty of, they're just sitting there and data science can be of help with that. And natural language processing so voice recognition and not just voice but also, of course, the language. The way people ... which words they use and so on to help monitor patients.

- Kirill Eremenko: So sounds like natural language processing is a crucial component of the work that you're doing. Could you tell us a bit more about that? What kind of tools or algorithms do you see the field of healthcare using in terms of natural language processing?
- Guillermo Cecchi: Okay, it's a very good question. So our approach is two prong. So one is the typical data science approach. So you give me data from like I said, these patients that not taking levodopa to improve their symptoms of Parkinson's. The same patients they are not on the drug and I will try to do a data-driven approach and create features and see what features are relevant and what features are not. That's one approach.
- Guillermo Cecchi: The other approach and I think this is interesting, perhaps more interesting. We tried to work with psychiatrists and neurologists to understand what are the concepts, what are the ideas, in some cases are explicit and in some cases are intuitive, that they use to judge a condition. To do a diagnosis, for instance.
- Guillermo Cecchi: And then we try to turn those into algorithms. So, let me tell you one example that is being very useful, that we developed for one condition but now it's something that we use across the board. At some point, we realized that, for understanding psychosis and that includes schizophrenia and mania, because of bipolar disorders, a very important concept is something that in the field is called flight of ideas. And the name, it's intuitive.
- Guillermo Cecchi: So, these patients they tend to be, for instance, I'm talking about a certain subject all of a sudden, boom,

they change the subject. Sort of non sequiturs and they're talking about something else. And that it's an essential component of the definition of what psychotic state is.

Guillermo Cecchi: So what we did was to use a technique that is while establishing natural language processing, there is semantic embedding that allows you to define vectors as the semantic content of the word. So, the word is represented by a vector and if you want to know how similar two words are instead of going to a thesaurus and see whether they are part of the same definition, you compare the vectors and the vectors are close, or the semantic content of the word is similar. And this is based on having computer statistics of co-occurrence. So, the whole idea is if the two words tend to occur together, for instance in passages, their semantic content is similar.

Guillermo Cecchi: And so we took that, which is used on a daily basis in NLP. And we created a measure of flights of ideas by saying okay so if in this phrase the average semantic context is one and the next phrase semantic content is very different because we can measure it, then that's an indication of flight of ideas. So based on that we were able to show that it's possible to predict onset of psychosis in patients who have sub-threshold symptoms. They are called clinical high-risk cohorts.

Guillermo Cecchi: And that's the interesting aspect of it is that this is the result of a dialogue with the psychiatrist that they know that this is important. But they don't know how to turn that into something measurable that they can share with other psychiatrists evaluating other people.

And what we brought from our side is tools that are already available in NLP to create something new. I think that's a good example of how we go about doing things.

Kirill Eremenko: That's a fantastic example. I've heard of that approach before where words are represented by vectors and closeness means similarity of semantics. I can see now that maybe this is yeah this is actually what you mentioned. You take processes or practices that psychologists already have, and they, kind of more intuitive, they develop them over time. And now you're putting them into something more measurable, quantifiable, some algorithm that actually works.

Kirill Eremenko: I think it's a whole art. And to take something that's intuitive and works through experience, through knowledge. One of the definitions of intuition I've heard is that it is experience and knowledge that you have that you know how to use but you cannot just describe and verbalize.

Guillermo Cecchi: Yeah. Yeah.

Kirill Eremenko: And so basically you're, you're creating algorithms that turn intuition, or you're turning intuition into algorithm. I think that's the next stage for data science for artificial intelligence. How do we create machines that not just do like ... are pre-programmed to do certain classifications or clustering on its own. But actually take something that's intuitive, that you cannot even the person that does that, has that practice. They cannot explain it to you.

- Kirill Eremenko: How do you take that and put it into an algorithm? And then that opens a whole new world because all of a sudden now you can apply it to mass scale. You can help many more people. You can combine the intuitive experiences and knowledge of different practitioners and see what you can come up with. So that's a really cool example of how you guys do this.
- Kirill Eremenko: Very interesting. So was this a recent research development that you guys come up with?
- Guillermo Cecchi: It's something that we published years ago, more or less.
- Kirill Eremenko: Three years.
- Guillermo Cecchi: And, yeah, and then last year we published a follow up showing that we can do these across cohorts, and even the interview protocol. So that's part of what we're trying to do. Try to understand how universal these features are. So if you interview a patient with a certain protocol in the certain context can you find information that is independent of that. That someone can use it in a different context, in a different institution, in a different country, in a different language. That's part of the goals that we set for ourselves.
- Kirill Eremenko: Gotcha. Yeah, very interesting. I actually was reading a paper, not a paper, an article recently. Where they said that through machine learning ... I don't remember where it was, but this the article sort of. Through deep learning, computer vision, machine learning is gone into stage that, and this has to do with teenagers who might be suffering depression. Through technology,



like that, we are able to discern certain breakthroughs recently that have enabled us to diagnose somebody who's suffering from depression, at about an 80 something percent accuracy rate. Which is even better than parents can diagnose their own children, which is ridiculous. Which is crazy.

Guillermo Cecchi: Yes.

Kirill Eremenko: So it's, yeah it's very exciting. What I wanted to talk about next was some of the limitations. So there's something we chatted about just before the podcast. And you mentioned that one of the major problems that you are seeing is that there's not enough data. That right now there's paradoxically, there's more brain imaging data, which is very complex to obtain, there's more of it than speech sample.

Kirill Eremenko: Speech samples of people with mental illness or people that might do ... that these algorithms can learn from. And in order for you to turn these findings and these research papers into actual clinical tools, you need larger data sets. Can you tell us a little bit more about that? Why is this a problem? It's so surprising. Like speech samples seem to be so easy to obtain compared to brain imaging. Why are there less of them?

Guillermo Cecchi: Well, a couple of reasons. One is that until recently the researchers and clinicians in mental health didn't think about the possibility of using data-driven approaches or algorithmic approaches to study language production, to study interviews, dialogue, etc. And in some cases, those interviews are recorded but are recorded, or were recorded with the purpose of

keeping a record. Not for instance with good audio quality. In some cases, we cannot use the audio because the recording quality is really bad.

Guillermo Cecchi: And the other component of it is, it's the issue of privacy. So everyone is concerned with the problem of privacy and the possibility of identifying someone who's suffering with a certain condition through their voice. And then there are all this centralized mechanisms in place, and now everyone is afraid of being sued. So, lawyers step in and they say, Well, between having data in our servers that might be potentially identifiable and not having that data, not having is safer.

Guillermo Cecchi: It's a problem but at the same time, we have the technology to solve this problem. So everyone does banking through their computers. And we are not afraid of being hacked although you know we are hacked and robbed of our savings. So, it's a problem that can be solved. But I think collectively we have to come to an understanding that it's something that we need to do. We need to have the system in place so that issues of privacy are solved. And they are solvable. That's my interpretation.

Kirill Eremenko: Everything is you don't need ... You don't only need just the words, right? Like one thing is if you just needed the actual textbook and then the audios can be converted to transcripts transcribed and then you can analyze those.

Guillermo Cecchi: Right.

- Kirill Eremenko: You actually need the voice, right? You need the patterns [crosstalk 00:34:09]
- Guillermo Cecchi: You know the voice contains, and conveys huge amount of information that the written form only partially conveys. So, prosody is very important. The whole point of sarcasm is that there's a mismatch between the words you say and how you say them. Right?
- Kirill Eremenko: Yeah.
- Guillermo Cecchi: There's a huge amount of emotional content that is conveyed primarily through prosody, and not necessarily through all the word forms. But, of course, we work with the transcripts whenever we have access to the sound files. We see that there is huge amount of information that is there.
- Guillermo Cecchi: And this is, it's important because it's part of the, as we were saying, of the intuition of someone who's doing a diagnosis of helping someone through [inaudible 00:35:25] the entire degrees of freedom of flexibility in language involved imply the acoustic-prosodic component. Right there is very important. So.
- Kirill Eremenko: Gotcha. Okay. All right. Actually out of curiosity, so two questions. What tools do you use ... what approaches do you use to extract that information about the voice? So we talked a little bit about how you use.[crosstalk 00:36:09]
- Kirill Eremenko: Sorry?
- Guillermo Cecchi: Yeah, so I'm actually, The first step is to build upon decades of research on speech recognition. So, this is a

very well established field. And so, we are using some of the basic low-level features that have been known to be very important to recognize voice. To identifying, for instance, vowels and discriminate vowels. And concepts such as the size of the space that your vowels span is important to discriminate. Different people to discriminate younger versus older, men versus women. So we build upon like I said basics of knowledge in terms of what are the basic features that are relevant for our processing of voice, in particular, the human voice.

- Guillermo Cecchi: You know, there are very well established tools to estimate features related to prosody at the level of pulses, duration of words. And, again, those are, most of those algorithms are readily available for anyone to use. They're in Python so you can be playing with that in very little time by just download packages. Start playing with it and get the signal related to, for instance, the pause distribution of someone who's speaking.
- Kirill Eremenko: Okay, and that was my next question. Do you use Python for all this research?
- Guillermo Cecchi: Well, for the most part, yes. For the most part, it's Python. It has really become the lingua franca of computer science.
- Kirill Eremenko: Okay. So, what role do you see deep learning playing in the space of voice recognition in the in the work that you're doing? Because some of the models as I understand it can be done absolutely without deep

learning. What about deep learning? Is there room for deep learning type of feature?

Guillermo Cecchi: Yes, Yes, yes. And we use it for some specific applications. Deep learning, or at least deeper neural networks as a way of incorporating, in particular, the idea, that meaning that we were talking about, what is the meaning of the word. The meaning of the word is only partially determined by the word form.

Guillermo Cecchi: So the context is important and in language when you're conveying meaning, that meaning it's distributed between the level of the word, the level of the sentence, the level of the paragraph. And so this idea of incorporating simultaneously different levels of abstraction sits very well on what you can do with deep learning.

Guillermo Cecchi: So, yeah from a purely data-driven approach it's always important if you have enough data to train the deep network. It's typically beneficial. But from a more conceptual point of view, it's particularly important when we are talking about something that is inherently multi-layer as language.

Kirill Eremenko: Alright. Okay, thank you. I think that's an important comment because deep learning is ... A lot of people are learning deep learning. And there's lots of different types of natural language processing so it's good to know that all of them kind of have room in this space in healthcare.

Kirill Eremenko: I want to ask you another question. I think you would be probably the best person to ask this. What are your thoughts on Neuralink? You know that whole notion

that Elon Musk is coming up with where they'll be connecting certain technologies directly to the brain. They'll be like a brain to technology interface, something like that. And like that will expand our .... well, hopefully, that'll expand memories and how fast we think and things like that. Do you have an opinion about Neuralink?

Guillermo Cecchi: Oh yes, of course.

Kirill Eremenko: Would love to hear like from somebody who's doing research.

Guillermo Cecchi: I have to disclose that a very good friend of mine is a director of that enterprise.

Kirill Eremenko: Oh wow.

Guillermo Cecchi: We met in ... so we were together in grad school in the same lab. So my opinion, well we can talk about this forever. For the time being, we don't really have a good way, or any way, actually, to decode information from the brain to transmit to another brain or a machine. Other than language.

Guillermo Cecchi: So, and the exception to this will be of course, people who have lost limbs, or lost movement. And then, you can use this approach to control devices or to control the robotic arm or whatever.

Guillermo Cecchi: So, for the time being ... first of all, we don't have a theory of the brain. I think we all agree on that. We are searching for a theory of the brain. So, if you give me all the information and all the spikes and all the snappy weights and everything at millisecond or some rate from the brain, I don't know what to do.

- Guillermo Cecchi: I can do a few things. But can I decode the brain of a writer and transcribe what the writer is thinking into a book? No, we cannot do that. Evolution found a very good way of doing it through language. And through some other mechanisms of communication that include body movement, facial movement, pheromones, and a few other things.
- Guillermo Cecchi: At the same time, I think that it can be done at the lower level. So we can have a direct connection conveying relatively simple signals. And that, I think, the value of that is pushing the boundaries, understanding what's possible, and forcing us to think along these lines. Along the lines of, well, what do we know about the brain? What information can we extract from the neurons and a few spikes?
- Guillermo Cecchi: So I think is extremely valuable from the point of view of pushing the boundary. In practical terms, I don't see much. Except for applications to people who have, again, lost limbs or lost movement.
- Guillermo Cecchi: At the same time, conceptually I think there is something that is coming. So we're converging between humans and machines. I spend most of my day surrounded by computers. One of the last things that I do before going to bed is check my cell phone. And then on my cell phone is my alarm and the first thing I do when I wake up is touch my cell phone.
- Guillermo Cecchi: Our connection with electronic media it's ever stronger. And, actually, computers in a way are mimicking our nervous system. So, it's digital and it's electronic. So that conversion is happening. Now we



are still very far. And I don't expect any clear advances from this approach. Except understanding where are the limitations.

Kirill Eremenko: Do you think it's good or bad that we are so connected with technology these days? Because if we think about it like you said, the only way we can express what is happening in our brains is through our senses or like actually like spoken language or touch and things like that. So, essentially, our mobile phones already an extension of us.

Guillermo Cecchi: Exactly.

Kirill Eremenko: Before actually, like, say, 10, 20 years ago, I didn't walk around on this., just like walking to school or whatever, I didn't carry all of the dictionaries and all of the world's knowledge with me in my pocket. And so if somebody asked me a question, I don't know, I don't know. But now I just open my phone and I look it up on Google and there's the answer.

Kirill Eremenko: So, essentially, it is an extension of our minds. Is just a very slow interface. The typing, your fingers, the looking at the answer of your eyes, the processing of it, is a very slow interface, but it's already an extension.

Kirill Eremenko: What Neuralink is proposing is to create that interface that's going to make it really fast that is directly linked to brain. And, as you mentioned, we don't have answers how to do that yet. But in general philosophically, is this whole idea of extending ourselves into machines, is it good or is it bad? Are there any dangers associated with that from your standpoint, from your professional point of view?

- Guillermo Cecchi: I wouldn't say professional, just personal. Because this goes beyond any particular profession. Is evolution good or bad? I think that this is not good or bad, this is inevitable.
- Guillermo Cecchi: And, we didn't choose to become human when we were primates. We are not choosing this. This is happening in spite of us. I mean there is no evil or for that matter good genius orchestrating this. It's part of evolution. We started with the cave. And the cave was an extension of the body or to do so far mother as a protection. And then we use tools as the extension of our hands.
- Guillermo Cecchi: This is an extension of our central nervous system and it's inevitable. It's not good or bad. There's always this question, every time there is a new technology. Is it good or bad? I always like to remind everyone that Socrates for people was against the written word. Because he thought that it softened the brain. Because until then, Homer could recite two entire sagas out of memory.
- Guillermo Cecchi: Now no one can do that because we just go out and read. Until recently we read the book. Now, you know, maybe, I don't know what we're gonna do. We have it in our computer and we don't read it.
- Guillermo Cecchi: So, every change in technology brings about these changes that some people want to say they're good or bad. They're inevitable. We can minimize the bad side of things, but we cannot ... we have to admit that this is happening. You cannot really stop it. Again it's like evolution. No one is pushing it.

- Guillermo Cecchi: We like it. The fact is that it's so close to us intuitively. Because it's an extension of our central nervous system. So it's very natural. It's like you give me a third arm, and after a while, I'll start to use it. So this is the same thing. It's there you're going to use it, and you want more. Because it's part of ... it's really very close to us in many ways.
- Kirill Eremenko: Okay, okay. I see your point. I've never thought of it as evolution. I always thought of it as like something different. Because evolution usually takes longer times. But nevertheless, I want to ask you then this, don't you think is happening a bit too fast? That we with all those other examples, whether it's speech, written text, tools in our hands, we had time to adapt to them and they weren't as life changing for us.
- Kirill Eremenko: Whereas now the amount of technological progress that's happened in the past 10 years is more than is how this happened, ever in the human history. And what we're seeing is things are starting to pop up that are affecting our 2 million-year-old brain in ways that we would never expect. For instance, the whole notion of social media. When you get a reply on Facebook or on WhatsApp or or whatever else, Instagram, you look at it and that causes certain hormones, I think endorphins to produce in your brain.
- Guillermo Cecchi: Yup.
- Kirill Eremenko: And that's actually the same process happening in your brain as when you're gambling, or when you're drinking alcohol. But we have age restrictions on both of those. Like alcohol and gambling, there's certain

ages, until which can't do it. And there's nothing like that on social media.

Kirill Eremenko: So children are getting affected by that. They're getting addicted to social media. Simply because it's only been introduced in the past couple years. And it's been so quick that legislation hasn't been able to keep up to speed. What are your thoughts on that?

Guillermo Cecchi: I'm not advocating the use of social media. I don't use social media. For a while, I was using the book tablet like Nook. I was using Nook.

Kirill Eremenko: Oh yeah, the one that blocks [crosstalk 00:53:12]

Guillermo Cecchi: Was using Nook to read novels and even that I dropped. And I went back to the old fashioned paper book. I think they are clearly not good in general. But at the same time, you know, you are saying changes are occurring very fast.

Guillermo Cecchi: When we invented bronze the poor people who had not bronze were wiped out. We killed them. And then when we invented iron we killed those who had bronze. And then when we invented the written word and we were able to create organized states, organized around books and read them principles, those states wiped out the older tribes and civilizations.

Guillermo Cecchi: So this happened many times before. Rapid technology is not new. And also let me remind you that we invented nuclear bombs a few decades ago. We were able to minimize. It could have been much worse. So, of course this is very, it's dramatic and it's fast. But it's not anything new. It is part of human history.

- Kirill Eremenko: Okay, very interesting point of view. Thank you. Thank you so much, Guillermo. On that note, we're actually coming to the end of the podcast. Thank you so much for coming on the show and sharing all your insights with us and all your opinions.
- Kirill Eremenko: Before I let you go what's the best way for our listeners to contact you? Maybe if somebody is interested in your research. You're in New York if maybe somebody wants to catch up and talk more about potentially collaborating. Or even just follow your work and career, what are some of the best ways to do that?
- Guillermo Cecchi: Email for the time being. So, just reach out and I'll ... like everyone, I'm overwhelmed with things to do and email etc. But I will try to answer.
- Kirill Eremenko: Alright, sounds good. So is it okay for us to share your email in the show notes?
- Guillermo Cecchi: Yes, yes.
- Kirill Eremenko: Okay, sounds good. And what about LinkedIn? Is that a place where people can connect with you as well?
- Guillermo Cecchi: Yeah. You know, I don't ... I try to minimize that with using email. So email if possible. So I don't get too many distractions.
- Kirill Eremenko: All right, like, yeah, like you said.
- Guillermo Cecchi: Yeah, it's, yeah.
- Kirill Eremenko: Gotcha. Okay. And one final question for you today. What's a book that you can recommend to our listeners that can potentially impact them or their careers?

- Guillermo Cecchi: Well I'm not doing anything related to what's in the book that I will mention. But it had a tremendous impact in my outlook, scientific outlook. And it's a book by Stephen Jay Gould. So anything that Stephen Jay Gould wrote was fantastic. He was an incredible writer.
- Guillermo Cecchi: But this book is about discovery of the fossil records of the pre-cambrian era where there was an explosion in forms of life. And the way he describes the discovery and weaves the theory of evolution around it and the task of the scientist, it's just phenomenal. I think it's called It's a Wonderful Life. It's about the Burgess Shale in Canada.
- Guillermo Cecchi: But if you can have any book by Stephen Jay Gould. He's probably one of the best writers of popular science. He was a real scientist, and it was a wonderful writer so. That's my recommendation.
- Kirill Eremenko: I think I found it online. It's called Wonderful Life: The Burgess Shale and the Nature of History by Stephen Jay Gould.
- Guillermo Cecchi: Alright.
- Kirill Eremenko: Awesome. Okay, well, on that note, once again, Guillermo thank you so much for coming on the show and sharing your insights and telling us about the brain, how data [inaudible 00:58:29] healthcare. Thank you very much.
- Guillermo Cecchi: Okay. Bye bye.
- Kirill Eremenko: So there you have it ladies and gentlemen. That was Dr. Guillermo Cecchi from IBM Research. I hope you

got some valuable takeaways from this podcast. It was very interesting to see how research has progressed and how much insights and data and information about patients can be extracted simply from voice samples. From the tonalities of their voices and the words that they're saying. It can actually help people get better. So it's really cool to see data science applied to help people, to improve people's lives and especially how people with mental disorders.

Kirill Eremenko: My personal favorite part of this podcast was probably when Dr. Cecchi was talking about how their approach of presenting words as vectors is one of the ways that they can accomplish transferring intuitive knowledge of professionals. Something that they use, that people and doctors in this space, that they have this knowledge they use, but they cannot verbalize. They cannot just explain how it is, how that knowledge is actually helpful for them in terms of diagnosing patients well.

Kirill Eremenko: This approach that Dr. Cecchi described is taking that intuitive knowledge and putting it into an algorithm. For me that's a next level of technology or of data science, machine learning, artificial intelligence. When you can actually encode not just simple rule based logic, but intuitive knowledge that is so hard to verbalize and communicate even between people. And by putting it into an algorithm that opens up huge new possibilities.

Kirill Eremenko: On that note, if you enjoy this podcast and you'd like to get in touch with Dr. Cecchi you can find all of the details including his email, and other things that we





mentioned on the podcast, such as additional materials and the transcript for this episode at [www.superdatascience.com/241](http://www.superdatascience.com/241). That's [superdatascience.com/241](http://superdatascience.com/241).

Kirill Eremenko: So make sure to reach out to Dr. Cecchi if this is a field that interests you, or you have some of your own ideas and thoughts on this topic. On that note, thank you so much for being here today. I look forward to seeing you next time and until then, happy analyzing.