Demetri Kofinas: What's up everybody? Welcome to another episode of Hidden Forces with me, Demetri Kofinas. Today we speak with Doctor Eric Schadt. Doctor Schadt is founding director of the Icahn Institute for Genomics and Multiscale Biology and Chair of the Department of Genetics and Genomic Sciences at the Icahn School of Medicine [00:00:30] at Mount Sinai. During the course of his 20-year career, Doctor Schadt has built genetics and systems biology groups at Merck, the Computational Biology Group at Rosetta, served as co-founder of Sage Bionetworks, as chief science officer of Pacific Biosciences, and now as founder of Sema4. He had published more than 300 peer reviewed papers in leading scientific journals and had contributed to a number of discoveries relating to the genetic basis for common, human diseases [00:01:00] such as diabetes, obesity, and Alzheimer's.

In this episode, we explore the information technology of biology. DNA: the world of genomics, where big data looms large. We begin by mapping the territory of the human genome, exploring the pathways of disease, understanding the ways in which complex genetic combinations express themselves as phenotypes like height, bone structure, intelligence, personality. [00:01:30] How are these traits coded for? What are the instructions our body uses to repair a damaged cell, to grow new arteries, regulate our appetite, or start us down the path of puberty? What happens when these instructions are damaged? How can the smallest difference in the order of life's code make all the difference for our success, our happiness, and even our very survival?

Fifty years have passed between the discovery of the double helix, and the mapping of the first human genome. [00:02:00] What progress have we made in the fifty years since? How has our ability to sequence new genomes, our genome, created a paradigm shift for the future of humanity? What is the role of big data and artificial intelligence in finding the correlations to treat malignancies, prevent disease, and improve the fate of our children? What is the promise? What are the perils? What stands in the way of us and this incredible, superhuman future?

[00:02:30] As always, you can gain access to reading lists put together by me ahead of every episode by visiting the show's website at HiddenForces.io. Lastly, if you are listening to this show on iTunes or Android, make sure to subscribe. If you like the show, write us a review, and if you want a sneak peek into how the sausage is made, of for special storylines, to open up pictures and questions, then like us on Facebook and follow us on Twitter and [00:03:00] Instagram at @hiddenforcespod. And now, let's get right to this week's conversation.

So, Doctor Eric Schadt, welcome to Hidden Forces.

Eric Schadt: Thanks, thanks for having me.

Demetri Kofinas: It's wonderful having you on. Actually, our last episode of Hidden Forces was on the qualitative universe, kind of looking at the limitations of data, and what the phenomenological world of interpretation and intuition and everything, but here is the exact opposite. We're going to look at the promise, the [00:03:30] world of data of big data,

specifically in so far as, genomics and date science is concerned. I think, before we start, I want to let our audience know, now, you are a pure mathematician.

Eric Schadt: Correct.

Demetri Kofinas: Which is awesome and fascinating and many of our audience will appreciate that from Episode 11 with Ray Monk where we talked about philosophical mathematics. You also have a PhD in biomathematics, which is something like computational mathematics.

Eric Schadt: Yeah, applied mathematics.

Demetri Kofinas: Applied mathematics within the context of biology.

Eric Schadt: [00:04:00] Yes.

Demetri Kofinas: Is that fair to just say that, that is a way of looking at the human organism, or organisms, or organic life in general as information processors?

Eric Schadt: Yeah, it's looking at living systems, more generally and the myriad of data that makes a living system live. And how are all those interconnected parts, how do they function, how can we model that? How can we build predictive model simulations to learn more about the operation of those systems, so that we can better tune [00:04:30] them, think about it from the standpoint of treating diseases, preventing diseases, and so on?

Demetri Kofinas: Fascinating, if we have a time, I want to ask you some more questions about that, and some will be peppered in during the course of this interview. I think to begin with, I want to ask you this. What is genomics? For me, genomics is the study of the human genome, but can you give us a better, sort of definition. What is genomics?

Eric Schadt: Genomics is studying the DNA that we're all born with all living systems possess some amount of DNA that really serves as the [00:05:00] blueprint for everything that's going on in your system. When you start with a single egg and a sperm that gets fertilized, there's two copies of DNA, one from mom, one from dad. Those come together and actually define-

Demetri Kofinas: Those are the chromosomes.

Eric Schadt: Those are the chromosomes, and they define all of the genes. The proteins that get generated that service the machines that carry out little functions in the cells, that enable the cells to build up constellations of cells, and ultimately tissues, and ultimately organs. Ultimately, these [00:05:30] organs that communicate with one another, that ultimately give rise to the complex being that you see sitting in front of you.

Demetri Kofinas: So, then to go further on that, what is a gene? I think when most people think of genes, they think of them as sort of, single units that exist somewhere in the body. You hear a lot of people say, "What's the gene for height?" Or, "What's the gene for diabetes?", but obviously that misses the point, so let's talk a little bit about that. What is a gene and what is that look like?

Eric Schadt: It's [00:06:00] a little more complicated concept than it may seem. In humans, you're born with the billion bases of DNA. That's spread across 23 chromosomes. Think of that DNA as an encyclopedia. In that encyclopedia, you have major chapters or themes, so those would be the chromosomes and within each of those chapters, you have words that then string into sentences, that convey [00:06:30] meaning. That's what the genes are sort of doing.

What genes are, are these units within the DNA that gets transcribed by machinery in the cell to produce the proteins that carry out all of these functions our cells need to survive. That's the simplest way to think about. That was the view back in the old, dogma of biology, back when we thought DNA went to RNA went to proteins. We now know that some RNA, these are the [00:07:00] RNA is transcribed off the DNA as the template from which the protein gets made.

Demetri Kofinas: They're like a reader of the DNA?

Eric Schadt: Yeah, they're a reader. There are special molecules, proteins that attach to the DNA, and then actually attach where the genes are located, so they have this ability to recognize, which is pretty amazing, where they should attach, where the gene starts, and then, it actually transcribes a copy of that DNA in the form of what we call, RNA. Then, it's that RNA that's then lifted off [00:07:30] the DNA and it goes into the ribosomes, where they make the protein.

Demetri Kofinas: When things go wrong in that process, that's where you can get something like cancer, right?

Eric Schadt: Right. Diseases, then are caused by, again go back to the book analogy, consider a gene as a sentence that's conveying some important meaning function of what should be happening in a system, and just imagine that now you go in, and you change certain letters or words in the sentence. You have the potential to fundamentally alter the meaning of that sentence. The same is true [00:08:00] in DNA, if you get a mutation in the single letters that make up the gene, that can cause the protein that ultimately gets transcribed to have and aberrant function or it can destroy its function. As is the case of cancer, it amplifies its function to a point where it's very, not normal, the cells just start proliferating and can't control themselves. Exactly right, so under normal conditions, these are being transcribed, everything's working in harmony, but certain mutations can occur that disrupt the system.

Demetri Kofinas: When the body looks to repair itself, it [00:08:30] references its genome, it references its DNA. It looks to see the blueprints, to see what do I need to create here, then, it proceeds to create that. If it needs to repair-

Eric Schadt: Exactly.

Demetri Kofinas: So, cancer is a situation where the blueprints or the information, that the body has available to it, in its genes, in its DNA is compromised. When it proceeds to try to build organic tissue off of that model, it creates something that is malignant, which is the cancer.

Eric Schadt: Yeah.

Demetri Kofinas: Or [00:09:00] a tumor that's benign or whatever, some tissue that shouldn't be there.

Eric Schadt: Yeah, more or less that's correct. Certain changes, we have what call oncogenes, which are genes that is they're expressed, proteins that occur in too high a number, they cause the cell to grow without limitation. You don't want mutations in those genes to get too active because then cancer results. Then, you have tumor suppressor genes, which are genes that keep all of that regulatory information in check for the cell, to prevent it from growing out of control and those functions can be destroyed. [00:09:30] If you destroy the tumor suppressor, you allow the oncogenes to proliferate, and they cause things to go crazy.

Demetri Kofinas: Okay. There's a tremendous amount of complexity, obviously in the gene, and there's another thing that's interesting, which is that of course, different genes have different lengths and different amounts of information because they aren't one, single unit. They are comprised of different nucleotide base pairs.

Eric Schadt: Yeah, that's correct. In fact, again if you go back to sort of the sentence analogy for a gene, the way a gene gets transcribed, there are these sort [00:10:00] of fundamental units, called exons. You can equate those to words. Those are being transcribed, but the complexity of the system allows some exons to be removed, and some to be put in for under normal functions, so it's pretty amazing. As the RNA's being transcribed from the DNA template, under certain conditions, it can say, we don't want this word, "We don't want this exon in this sentence, so toss it," or, "We want to make sure this one's in there." So, there could be variations in the gene, [00:10:30] in the functioning of the protein that's needed to adapt to whatever conditions you're facing.

The complexities even at the gene level's pretty high, but it's important to note that complex diseases like cancer or diabetes or obesity or Alzheimer's, those are caused not by a single gene, typically, or a single problem, or a mutation in the gene. They're caused by a constellation of changes over many genes. Again, it's disruption, not at the sentence level, but at the chapter level. It's that, [00:11:00] sort of more network oriented, so think of each gene as communicating with each other on what level should I be expressed at, who do I

join forces with to make this machine work better or to stop this machine from functioning. It's this network that kind of goes out of control, not a single gene.

Demetri Kofinas: Exactly, the complexity. And the same is true also for certain traits like height. Of course, some mutations are beneficial, and so that's the way evolve and everything else. The same mechanism that causes cancer, in some way or another, is part of the evolutionary [00:11:30] process.

Eric Schadt: Absolutely.

Demetri Kofinas: I hope that sort of gave our audience a basic template. Now, that's get into the technology of genomics, which is where you reside. I guess, let me ask you first, what is the problem that Sema4, your company, is trying to resolve or solve?

Eric Schadt: The problem, Sema4 is trying to solve, address is how do we better understand the human genome and the changes [00:12:00] that occur in that genome, that either protect you from disease, enhance wellness, think cognitive functioning, all the favorable things, but that also can cause bad things, Alzheimer's, diabetes, cancer. What we note, is we have the ability now to sequence entire genomes in people for low cost at population scale. What we're missing is the high dimensional data around people to understand, what are those changes in DNA doing, what are they causing to happen. You need phenotypes, [00:12:30] so think of when you go to the doctor's office, your blood pressure, your glucose levels-

Demetri Kofinas: The phenotype is the way in which the genotype expresses itself-

Eric Schadt: Exactly.

Demetri Kofinas: And the way we experience it.

Eric Schadt: Perfect. That's exactly right, and if you think about the number of phenotypes that can exist, they're in the hundreds of thousands to millions, so we can consider, phenotypes could be the expression levels of a given protein in a certain part of your body. They're 20 to 30,000 different proteins, or variations of proteins that are going on, so [00:13:00] all the different combinations are pretty enormous. Think of all the epigenetic changes, these are chemical modifications that can be induced by the environment to change the DNA, those can occur by the millions.

Think of all the environmental exposures, the pathogens you're exposed to, viruses, bacteria, pollution levels, pollen counts. All of that information molds who you are and how your genome's expressed. If you want to get to a complete understanding of disease, you need to have all of that information built up around populations of individuals, and that's [00:13:30] Sema4's mission.

Demetri Kofinas: Just to clarify something, epigenetics, because that's become more of a buzzword in the last few years with more and more people are becoming aware of it. Is

epigenetics the study of the way in which genes turn on or off certain functions or characteristics based on accommodative environments?

Eric Schadt: I think that's close. Mutations are changes in DNA, are actually changing the fundamental letters in the sequence of the DNA, and then that can be passed on to your kids, and their kids, and [00:14:00] so on. It's a heritable unit. Epigenetic modifications, you're not changing the letter, itself. You're chemically modifying the letter to have a different kind of flavor.

Demetri Kofinas: That's fascinating.

Eric Schadt: That chemical modification can be induced by environmental exposures, like endocrine disrupters or common environmental toxin, a lot of supermarkets use it to spray off their fruit; that can induce chemical modifications to the DNA, you're exactly right, can turn on [00:14:30] and off genes in certain programs to wreak havoc on the system. It's an interesting kind of change that we've evolved with, to allow a more rapid adaption to the environment to tune what's going in the system. It's amazing.

Demetri Kofinas: Amazing. Is it also true that given sufficient generations, epigenetic changes become genetic?

Eric Schadt: Yeah, and that's been the big finding, this must have been, maybe seven or eight years ago, maybe even a decade. [00:15:00] One of the biggest findings to come up and my colleague, Doctor Sakina at Washington State, was one of the first to do this. They found that endocrine disrupters, so this is a common environmental toxin, that induces these chemical modifications to the DNA can actually be passed on to future generations. What they were able to show is those things that were environmentally induced where then passed on to future generations, and the same sort of molecular [00:15:30] change that induced in the primary affected, was passed onto generations. That was never thought to have occurred.

In fact, we used to joke about the Lamarckian evolution, that remember, the giraffe's neck is long because it reaches for the trees, and everybody would laugh at that, that's not how evolution works. What we're seeing for the first time is absolutely, environmental changes can induce these modifications that can be passed on, and they will have an evolutionary effect.

Demetri Kofinas: Lamarck had the theory that was a little [00:16:00] earlier, but also competing with Darwin.

Eric Schadt: Correct.

Demetri Kofinas: I mean, he was around the same time. What about telomeres because that's something I've also learned about in recent years. I think that's even more recent in some ways in the popular culture to the extent that it's popular is telomeres. Tell

us a little bit about that because and I'm asking about these questions why, for two reasons for our audience to know. One is, I want to highlight the complexity because I want people to appreciate that there is sort of, the what we would like, and then the reality of getting there, which is extremely difficult.

The second point [00:16:30] it sort of related to this, which is as we get deeper into what you do, you're dealing with data, you're dealing with looking at genetics, epigenetics, things like telomeres. Every piece of data that you're coming in, you're trying to match that to phenotypic expressions. How it expresses whether as traits, characteristics or disease, and you're trying to find correlations. I just want to express that point, that you're collecting massive amounts of data and we don't even know fully how all these systems operate with each other.

Eric Schadt: Yeah, and that's a good [00:17:00] point. Things like epigenetic modifications, telomeres, they all sort of relate to how is DNA managed in the system in the cell. So, DNA, think it's three billion letters long, so it's pretty long. The only way to efficiently store that is to kind of wrap it, as to kind of compress it down, wrap it around proteins that we called histones to make these chromosomes, and then the chromosomes get capped with these telomeres, which sort of ensure the integrity of the chromosome.

What's come [00:17:30] out in the last decade is that the telomere length is sort of intimately intertwined with aging. The hypothesis is that as you age, damage is occurring, telomere lengths decrease, the stability of the chromosomes are then in jeopardy, and as time goes on, it sort of loses its coherence and the cells ultimately function at a suboptimal level, suboptimal enough to where you die in old [00:18:00] age.

Demetri Kofinas: It's like a protective of the gene, of the genome or whatever the correct term is.

Eric Schadt: It's protective of the chromosomes that contain all of the genes and that stretch of DNA, so you want healthy, long telomeres to maintain that kind of robustness, and again the aging process or toxins in the environment. Those are things that can degrade, not only mutate the DNA, but can shorten the telomeres and basically progress you; we now distinguish between your chronological age and your biological age. You can [00:18:30] be a 30 year old chronologically aged individual, but your molecular states look like your 40 or 50. How does that happen? That could be through your increased exposures to toxins that are mutating your DNA, shortening your telomeres, and so on, and causing sort of aberrant functions in the pathways that you need for normal functioning.

Demetri Kofinas: What you're also getting at, which is diseases like cancer are really the body's inability to keep up with the repairs with the damage. They aren't able to keep up.

Eric Schadt: Yeah-

Demetri Kofinas: [00:19:00] 'Cause we're developing cancer cells are the time in our bodies. The

Eric Schadt: That's a great point and it is this balance, this steady state that your system gets into, and can maintain the appropriate pushback on cells that may become aberrant, and so your immune system activates, and gets rid of those. If you destroy that imbalance, you suppress those systems, you're exactly right. It now provides an opportunity to proliferate those cells.

That's one of the most exciting components of the immunotherapies [00:19:30] that you're hearing in the news today. What is that doing? That's amping up your system's immune system to attack these cells that are growing out of control, that have a specific signature, the body can sense that they're cancer-like, they shouldn't be there. What the immunotherapies do now, is they boost your immune system's ability to both recognize those cells and destroy them.

Demetri Kofinas: We're aiding intelligently the immune system to recognize-

Eric Schadt: Exactly.

Demetri Kofinas: And attack.

Eric Schadt: Yeah.

Demetri Kofinas: That leads us into a really good question. So, where are we now in terms [00:20:00] of data science, big data, and genomics, and pairing all those together towards creating treatment modalities? Where are we at this moment? Before we get into where we could go.

Eric Schadt: Where we're at, at the moment is we've dumped a tremendous amount of money into sequencing many genomes across many populations, disease groups across a whole array of diseases, different populations, well individuals. The thing to notice, DNA sequencing technology [00:20:30] is the only technology known to human kind that is moved at super Moore's Law speed. Think fast or exponentially faster than semi-conductors and memory and so on. It's amazing, and that's moved with such amazing speed, where the cost has dropped so fast, that we have been able to literally sequence millions of individuals now. That's good. That gives us, [crosstalk 00:20:54].

Demetri Kofinas: Why is that? Is that correct, it's about five times.

Eric Schadt: Yeah, something on that order.

Demetri Kofinas: Why is that? What is [00:21:00] the reason for that?

Eric Schadt: That's a good question. I think it's just the emergent, the bringing together of many different types of technology that have been able to benefit the caring out

of molecular assays, like sequencing DNA or sequencing RNA, that I think that space has benefited from all of the nano-fabrication, nanotechnology, the semi-conductor industry, all of those technologies being brought together in a synergistic way.

Demetri Kofinas: What was [00:21:30] the cost? Like a hundred million ... Am I incorrect? Was it $100,000,000 in 2001 to sequence the human genome?

Eric Schadt: More on the order of a billion to three billion.

Demetri Kofinas: And when we hear the human genome, what we really mean is one person's genome, some lab rat.

Eric Schadt: Not only is it one person's genome, but it wasn't even the whole genome. We talk about the human genome has been sequenced, and at that time, there's probably 20 to 30% of the genome that was not-

Demetri Kofinas: That cost could've-

Eric Schadt: Could not be addressed by that technology.

Demetri Kofinas: And now it costs under a $1,000?

Eric Schadt: And now it costs, [00:22:00] yeah-

Demetri Kofinas: Just about a thousand.

Eric Schadt: On the order of a thousand.

Demetri Kofinas: That is pretty remarkable, right?

Eric Schadt: Very remarkable.

Demetri Kofinas: Because how many years is that 16 years?

Eric Schadt: The first paper was published in what, 2001?

Demetri Kofinas: Yeah. It's amazing.

Eric Schadt: It took it like 15 years.

Demetri Kofinas: The way I see it, there are two major bottlenecks, and it's not clear to me, which one is the bigger bottleneck out at the moment, or which one will be. I can guess, but one is the data, which is sequencing and acquiring, not just sequencing, that's an interesting point [00:22:30] as well. There's the sequencing, which is getting new data, and even better data, and then the other part is getting existing data, that's locked here, or protected over there. And then there is the software side, right?

Eric Schadt: Correct.

Demetri Kofinas: Where are we there and tell us a little bit about what you're doing with that?

Eric Schadt: I would say, first of all, while we've made amazing progress on the DNA side, the fundamental problem we're faced with today, in my opinion is building up [00:23:00] enough other types of data around individuals, where we have that genomic information, where we can link it together. Today, we simply don't have, well, all this money's gone into sequencing, many, many individuals, if you look at the studies that have been done. It was looking at obesity and diabetes and heart disease and cancer, but those individuals weren't sequenced exhaustively. They were sequenced, "Do you have diabetes or do you not have diabetes?", so really minimal amount of information was generated on those individuals, and not enough to really [00:23:30] get at the complexity of what's going on.

One of the fundamental problems we have today is building up a big enough information store around individuals to hook the DNA up to. Think of, so what data do I mean? When I mean molecular data, so think of metabolite levels, the vitamins you take. Think of protein levels, the machines in your cells that are doing all the work. Think of the epigenetic changes, and so on. We don't have exhaustive amounts of data built up-

Demetri Kofinas: Are there no standards? Or were there [00:24:00] no standards?

Eric Schadt: I think it's the technology's moving so fast, that it's only been within the last few years, that we've had the capability to generate that information. That's molecular, now think of, we want to get more to the physiologic. What do we want to do? We want to hook up all this molecular information to the physiologic information, so that we can understand how they interrelate and how we can push on certain parts of the molecular system to impact the physiology, to prevent disease, to enhance functioning, and so on. I think on the physiology side, we've been pretty limited.

[00:24:30] Think about when you would go to the doctor's office 20 years ago, it's a doctor with a stethoscope listening to your heart. Today, we have a myriad of molecular tests that are run, we have imaging technologies, we have advanced devices that we can wear on us that are monitoring us longitudinally. We're in this revolution, as well, of being able to generate really high dimensional longitudinal data on our physiological states, and that's the kind of data that we need to build in mass to compliment that big [00:25:00] DNA that's been generated.

Demetri Kofinas: So, if anything, it's not the ... Acquiring the data is less of the issue. A bigger part of the issue is organizing an and sorting through what's out there. Also, from what I understand, informed consent's a real issue, so a lot of this legacy data, what percentage of that is something where you could go back, and contact the patient, and get an update, and update their records?

Eric Schadt: Very minimal. In fact last year, we published one of the largest genome studies, ever. We looked at 600,000 [00:25:30] genomes, a paper we published in Nature Biotechnology, and the idea was to look over these 600,000 genomes, that we went to all, everybody on the planet was generating DNA data, and said, "Hey, can you be part of our study?" The study was, can we identify individuals, who harbor very deleterious changed in their DNA, that should have cost catastrophic illness when they were a kid.

They should have been dead before adulthood, but they're in their 40s and 50s, and they're fine, and they've never manifested the disease. The medical [00:26:00] text books says, "Should be dead," but here they are, alive, never had the disease. However, they're buffering the disease, however the nature allowed them to counteract it, is the therapeutic. So, it's a way to identify, immediately, a therapeutic that could benefit whole disease groups.

We do this screen over 600,000 genomes for this special individuals, that we call unexpected heroes and identified around 13 candidates. We weren't able to gene and recontact one of them.

Demetri Kofinas: Legally.

Eric Schadt: Legally, because the way the consent, like all of the different [00:26:30] studies that were done, the types of consents they had, just didn't allow for that. And in cases where the consents were there, there were serious concerns about privacy, about going back to somebody and relaying this information, where the institution didn't want to take on the risk of being sued.

Demetri Kofinas: You're bringing up another huge point, which is privacy, and regulations and all that. Before we go to there, because I do want to ask you about that. How many human genomes have we sequenced?

Eric Schadt: So, there are many different measures of what it means to sequence a human genome. [00:27:00] It's complicated. We have what we call whole genome sequencing, which is probably what you're thinking when you say, you get your whole genome sequenced. That's literally the attempt, to sequence your entire genome. You don't necessarily have sequence your entire genome to get the most actionable stuff. We can sequence just the genes.

Demetri Kofinas: Because when you say sequence the whole, human genome, do you include the redundant parts in that?

Eric Schadt: Absolutely, so those of us who work in information theory and network, network reconstructions, we [00:27:30] want it all, because even though the genes, the protein coding parts of the genome comprise maybe three percent of the genome. So, what's the other 97% doing? Well, they're coding for non-protein coding genes, they're coding for regulatory regions. Remember when I talked about the proteins that bind to transcribe the gene off the template. Those are regions are pretty important.

Demetri Kofinas: The RNA.

Eric Schadt: They have functions, that if you really want to understand how this network is put together, you need to know it. For medically [00:28:00] actionable, for hooking up what changes in DNA have the most profound impact on disease, the protein coding component of the genome is pretty informative, and it's ... Whereas a whole genome, sequenced that clinical death, is going to cost you 1500 to 2500 dollars on an exon, which is the protein coding part of the genome.

We would sequence about three or four hundred dollars, so it's still a big enough price difference, where the bang for the buck will be done there. On the whole genome sequencing [00:28:30] side, we've maybe sequenced on the order of a hundred to two hundred thousand genomes. On the whole exon side, easily now into the low millions. On the genotyping side, which covers all the common variation that would exist in the population, tens of millions.

Demetri Kofinas: When I hear the number 10,000,000, which is sort of the target that has been set for being the sufficient number [00:29:00] of genomes sequenced to get us to a place where we can make important, meaningful insights. Where did that number come from? Where are we in terms of that number and ...

Eric Schadt: So, it's a good question, and I think the 10,000,000 number, again think of it in terms of, if all we have are 10,000,000 genome sequenced, without all the data built up around those individuals, the data's not useless, but it's far, far, far less useful. If you have high dimensional data built up around those 10,000,000 [00:29:30] individuals, plus their genome. That's what's going to transform our understanding of disease, and why is that? Think of something like cancer.

At first, let's go to information theory. Think of the Googles and the Amazons, look at the papers they publish on deep learning, machine learning algorithms. What does it take to derive from high dimensional feature data? What sample size do you need to really have these methods work their magic, and pull out the most predictive features that are amazing? It's on the order of ten [00:30:00] to twenty thousand. Sample sizes of ten to twenty thousand.

Now, think of cancer hundreds of different types of cancer, in a given type of cancer there'll be many different subtypes of cancer. If you want to get 10 to 20,000 people who are homogenous enough with respect to a certain type or subtype of culture, you need 10,000,000 people, the sample to get to that number. Now think about you want to go across the hundreds of individuals. You don't want to just target, recruiting those 10, to 20,000 in that one type of [00:30:30] cancer because it's only gonna be good for that one type of cancer. You want the 10,000,000 where now you can sample and start covering this entire space of disease.

Demetri Kofinas: It could be possible once you get the 10,000,000 disciples, we really need to get to a 100,000,000, then eventually we just want all the data.

Eric Schadt: Everybody.

Demetri Kofinas: That's fascinating. You mentioned the software, and machine learning and artificial intelligence is very important in this process, right? The one side is the data which we discussed, there are many complications to getting it, there is regulation. Before we [00:31:00] get to the sauce, what are the regulatory hurdles that you guys are dealing with? What are you doing also on the business development side to acquire this data where it exists?

One thing that I was thinking about when you were speaking is how useful will it be to be able to track these people over their lifetime? That gets the point of informed consent in, being able not just to have the really great data now that covers all the bases you are describing, but being able to go back to the well and update the phenotypes.

Eric Schadt: On the first point, [00:31:30] DNA, the sequencing of DNA and conveying to individuals risk information or even paternity, the parents you think are they your parents, your disease risk and so on. That's heavily regulated. It's heavily regulated and the state of New York is the most heavily regulated of all the states in the US, and the New York department of health requires pre-market approval of any genomic test on any citizen of New York that's gonna [00:32:00] convey medical information back to the individual.

You have to go through the same kind of FDA regulatory hurdles that you would have to if you were making a drug or some other device. It's something that definitely exists, it is a high hurdle. It makes some sense in that DNA doe identify who you are. You don't want people basing decisions on their healthcare, their choice of drugs for therapeutic, for treatments [00:32:30] based on information that's not correct or we don't know the confidence of it and so on. There is some rationale for why that's true. That's why one of the strategies for Sema4 in engaging individuals in this information is built up around testing of individuals in a clinically relevant context. For example, we carry out on, think of a pregnancy journey for mom where she wants to become pregnant.

The first test she might have is the carrier screening test, which is gonna [00:33:00] say do you carry mutations that you may pass onto your child that could cause really bad disease? You should be aware of that. Especially if your reproductive partner carries the same kind of bad mutation, then you are at super high risk of having a baby with a disease that again you should be counseled around.

Demetri Kofinas: This is known as genetic counseling when people-

Eric Schadt: Genetic counseling, you are counseling the individual around that, and then once the baby, once the fetus is in play in the [00:33:30] woman, you have non evasive prenatal testing, and that's to look at are there any really bad gross DNA changes that are gonna lead to catastrophic disease? Then the baby is born and you have newborn screening to say, can we aid in the diagnosis of disease that would facilitate more rapid treatments so you have a better outcome of the kid. We have a very natural point of

engagement of individuals who are going through say a pregnancy journey, to engage them in that testing, counsel them around the results, but then also [00:34:00] educate them about the utility of that data beyond that specific journey.

The idea is to form longitudinal longer term relationship with that individual to track everything that's going on in their lives over time. As the rate of knowledge is changing, the interpretation, the genomes change, we can update them on we told you this two years ago, but the current state of knowledge says this, and it may change guidance around the risk of disease or the risk of their child developing a disease, [00:34:30] and how you might intervene and prevent that disease from occurring.

That's our strategy is on engaging patients longitudinally in their journey and sort of maintaining this relationship with them to guide them as the states of knowledge change, as the amount of data we can acquire around them changes, and as our models become better.

Demetri Kofinas: Because the acquisition of data is also, the method by which you acquire that data directly, new data is a product as well of Sema4.

Eric Schadt: Yeah. It's a good point on whether that's a product or not. Our [00:35:00] sort of position is that when a patient has data generated on them say within a health system, they to go the physician's office and they have data generated on them when we generate DNA sequencing on them for these tests. My view it the patient should own that data. That's their data. What we want to do is enable patients to take control of their own data, help them manage that in a way that they don't have to have a PhD in computer science to do it, with tools that are easy to use, where they can [00:35:30] dictate where do they want that data, who do they want the data shared with?

As they flip from health system to health system or they move can they share that information with their new physician? Basically just enable them to have that kind of control. Sure, we would like an non-exclusive sort of rights back to the data to help aggregate it, build better models that we can feedback into the diagnostic engine to make better interpretations of the information. That's the model we are after.

Demetri Kofinas: Listening to you talk, I don't know if I've ever encountered a more complex business [00:36:00] than the businesses in this field. You've got massive challenges to overcome for sure two if not three or more areas. Just business-wise, I wasn't even thinking about the regulatory side. You've got so many challenges in so many fronts. It's just amazing.

Eric Schadt: It's important to point out that we, Sema4 is a new company we are not a startup. The enterprise they built at the Mt Sinai Health system Icahn School of Medicine. We had this existing [00:36:30] business where we are performing these tests, and carrier screening and non-evasive prenatal testing and cancer genomics. Those are tests that we run during the course of somebody's care, and those are actually reimbursed.

We actually get reimbursed through payers, insurance companies on those tests for those individuals.

Again, one of the beauties of Sema4 is we are not having to buy the data, like some companies are having to do. We are actually generating the data in the course of somebody's care. We are getting reimbursed for [00:37:00] that. The investment, and you are right, there is a lot of complexity is building the big data store to manage all the information, to compute on it, to drive predictive models that can then go back and in form. That is an expensive game. That's not, to hire teams of data scientists.

Demetri Kofinas: You guys have hundreds of people working for you right?

Eric Schadt: Sema4, we started with about 300 people, but we feel like we are gonna need in the hundreds of data scientists if we really want to mine the data deeply enough to derive all the most meaningful insights. I use [00:37:30] Tesla, Tesla is an example ... They have near to 400 to 500 data scientists that are doing nothing but looking at the sensory data coming off the Tesla cars to improve the autopilot. What we are trying to do so is far more complicated than improving the autopilot.

Demetri Kofinas: Because it's more complex systems.

Eric Schadt: Because it's more complex systems. One could argue that you need to run a scale on this information side. You need to be at the hundreds. In the four to 500 that Tesla has today that's more than the sum total of all data scientists across all the health systems in the US. [00:38:00] Again, nobody is really thinking about the scale you have to run at to be successful or derive meaningful insights from all of that big data, and that takes big investment to do.

Demetri Kofinas: We are gonna get to that too, because that's so important, because at the end of the day you can have all the data in the world, but knowing what question to ask and what to look for is really the key. I would love to ask you about that.

Eric Schadt: That's a good point. I have to just say-

Demetri Kofinas: Go ahead.

Eric Schadt: When we were out pitching to raise money, I was sort of selling the, if we build this big data store they will come. [00:38:30] We have this business we are making money. One of the investors said back to me, and he goes, "Do you remember Ask Jeeves?" Immediately I remembered-

Demetri Kofinas: I remember that-

Eric Schadt: I remembered but you haven't of Ask Jeeves since you remembered. Ask Jeeves at one time had more data indexed on the internet than google. Had this big information store, but didn't have the right way to engage the information and query it and

provide meaningful, enough meaning back to motivate people to use this. [00:39:00] You are right, you need to have the right questions to ask.

Demetri Kofinas: Data is not gonna say anything to you, you have to ... What's also fascinating listening to you talk is really appreciating how brand new this whole field is, it's so coming together rapidly. There is so much opportunity for people and companies who are asking the right questions and have the right solutions, remarkable. Let's get a little deeper into the software. Tell me a little bit about that. In general, are you guys looking for correlations? [00:39:30] Let me understand a little bit better how that process works, to the extent that you can even talk about your software.

Eric Schadt: I think again we are generating lots of information on people. For each person we have a column in a big feature table. Each of the rows are hundreds of thousands of rows that are all the different variables we are measuring on that individual, whether it's changes in DNA, RNA levels, protein levels, metabolite levels, glucose levels, weight, blood [00:40:00] pressure, pollen count exposure. As many things as you can think. What do you want to do with those hundreds of rows in each column, or you want to link them to lots of columns, lots of individuals, because the cross-populations of individuals you are gonna be able to start understanding how do these variables relate to one another?

How correlated are these two variables with each other, how do they correlate with everything else? When you start thinking about how do you tackle that kind of problem you start thinking in terms of [00:40:30] networks. A network is simply a graphical model where each node, each unit in the network represents one of those hundreds of thousands of features and the edges and the networks, the connections between those nodes represent a degree of relatedness between those, how correlated are they through the population you mapped?

Demetri Kofinas: The visualization component.

Eric Schadt: It's visualization, but it also reflects, it's visualizing statistical model underneath. I don't want to get too off the rails here, but [00:41:00] you can view a network as a joint probability function, where what you are looking for is the probability that your inner state giving these hundreds of thousands of variables of information. That's a pretty complicated probability function. What we do is we come up with network based algorithms, statistical algorithms that try to decompose that very complicated function into easier units that we can actually tackle on the computer.

At the end of the day what we are trying to do is [00:41:30] say what's the interconnectedness what's the topology of all these variables that may inform on disease, and that topology helps inform us what are the subunits that are actually driving disease X? What's the different topology that may be driving disease Y? What's the components that enhance wellness? It's this big statistical model that you can then query to start asking questions about what happens if I move this node in a certain direction.

Demetri Kofinas: That's so exciting.

Eric Schadt: It's cool. Think [00:42:00] of it as a simulation ultimately, what we are doing is simulating the molecular state of the system on the computer and allowing biologists instead of doing the lab experiment of knocking a gene out in the mouth which takes months, and then phenotyping it. You can carry that out in seconds on the computer, by doing In Silico perturbations on this representative model of the system. That's the kind of goal we are trying to achieve.

Demetri Kofinas: That's actually something that other sciences have had for a long time.

Eric Schadt: Correct, yes.

Demetri Kofinas: In physics, we model the world [00:42:30] before we ever actually try to conduct experiments. Are you trying to do the same thing for medicine?

Eric Schadt: Perfect. It's absolutely right. When the Hadron Collider generates the petabytes of data it generates, the physicists, they don't go and de novo reconstruct all the laws of physics. They project that information onto existing models, and assess the degree to which the models worked and predicted what they thought, and didn't work, and when it doesn't work they are coming up with new theories, new hypothesis, [00:43:00] new ways of modeling. In biology, health or biomedical sciences absolutely has to evolve to that sort of harder quantitative science, like physics, climatology, even quantitative finance.

We have to start storing knowledge and understanding not in papers, but in statistical models that can be queried, they can be assessed for accuracy. They can be refined, they can be adapted, they can learn. That's the path we have to go. [00:43:30] That's what Sema4 is trying to push, one of the first waves of doing that.

Demetri Kofinas: Creating a really robust model that's predictive.

Eric Schadt: Creating [crosstalk 00:43:38] first incidence of a robust model that's predictive but is also adaptive and can learn as the data builds, as we see more of the outcomes, we can go back and refit the model, make adjustments, and that the model should just get better and better over time.

Demetri Kofinas: The same sort of bottlenecks, is it fair to say the same bottlenecks, impediments, whatever you may call it, on the software side [00:44:00] that you face are those that any data science company like google or Facebook is facing? It's really on the progression of artificial intelligence and machine learning?

Eric Schadt: Yeah. I think it's close. I think the difference between what google would do where their primary aim with all the information is more around classification. It's how can we predict what group do you look like, given all the features? Then what will you respond to if I sent you this ad? Whereas [00:44:30] what we are trying to do is learn the fundamental rules of biology. We are trying to get at mechanism, actually come up with models that describe the mechanistic underpinnings of the system that then can go on to

offer sort of the predictive components that we need to actually to do better treatments, diagnosis if disease.

Demetri Kofinas: Within that process, are you able to determine causation?

Eric Schadt: Yeah. In fact one of my focuses over the last 15 plus years [00:45:00] has been how do you go from correlation to causation? Most of the network reconstruction algorithms that people employ, just look at correlative information and there are many problems with trying to go from correlation to causation, because the way information can be structured, you can have multiple different models, multiple ways of relating variables that are statistically indistinguishable from one another.

No matter how much data you have you are not gonna be able [00:45:30] to resolve it. You need some kind of, like perturbation to break that symmetry and then observe what's the most likely causal relationship. What I had helped pioneer 15 or so years ago was leveraging DNA variation, so changes in your DNA as a perturbation source to understand how these hundreds of thousands of variables are causally connected.

Changes in DNA as we discussed in the beginning cause changes in the function of a protein. [00:46:00] The changes in the function of that protein cause a cascade. They cause changes in the way the cell functions, changes in the way the tissue functions. Changes in your physiological state, your risk of disease and so on. We can now monitor that cascade and employ these advanced algorithms that can consider DNA as a perturbation source, and through some sophisticated mathematics start making causal inferences from this correlative data. That was one of the, I think big advances in being able to go from [00:46:30] this high dimensional descriptive correlative to actually causal mechanistic models of disease.

Demetri Kofinas: What are some of the most interesting insights you've heard in the last few years in this process?

Eric Schadt: I think the biggest most fundamental one would be that diseases are caused by perturbations to networks, not single genes.

Demetri Kofinas: Changes.

Eric Schadt: Changes. This was something ... Think of a drug company. What do all the drug companies, what's their main aim? It's to find a [00:47:00] single disease gene protein to build a single small molecule, to target that, and thus correct the disease. What we've seen over the last decades is that has, there has been some successes but by and large that's fundamentally failed. The work we were carrying out showed why did that fail, is because it's not any single gene or defect that's typically causing these common diseases, it's sort of changes in the states of these networks, which are caused by many different genes, and changes [00:47:30] in those genes and impacts from the environment. If you want to-

Demetri Kofinas: The feedback of the medication too, that's gonna be an interesting angle.

Eric Schadt: Right, all of those elements that you can imagine that your medication, the viral pathogens you are exposed to, the bacterial pathogens you are exposed to, the pollution levels. All of that is having an impact on the-

Demetri Kofinas: And changing itself.

Eric Schadt: And changing itself. You need to be in this fluid dynamic modeling state to actually achieve that understanding. That's really a fundamental difference from the reductionist sort of biology that drove the last 100 [00:48:00] years, which says if the central dogma biologist was DNA to RNA to protein. I would say the new dogma is constellations of DNA combined with environment altering networks of molecular states that go onto change functions [crosstalk 00:48:16].

Demetri Kofinas: Embracing complexity, complexity theory.

Eric Schadt: Embracing complexity.

Demetri Kofinas: Yeah, which is something we see across a lot of disciplines. On this show we talk about a lot in the context of economics as well as sort of a way of understanding economy. The financial markets that, Newtonian, [00:48:30] new classical models of economics mess.

Eric Schadt: Perfect. I would say biology is definitely a latecomer to that complexity space.

Demetri Kofinas: And arguably the most complex.

Eric Schadt: And arguably the most complex, absolutely.

Demetri Kofinas: What's going on with wearable technology? I know that apple has some type of API that it shares with companies like you. What is the opportunity there in terms of being able to get real-time data from people?

Eric Schadt: I think it's huge, because if you think about ... [00:49:00] There is lots of sophistication in the doctor's office. You go to the doctor's office and they now have lots of equipment and can run lots of test, so what's the problem? The problem is how often are you in that doctor's office?

Demetri Kofinas: Hopefully not often.

Eric Schadt: Hopefully not often. For the normal person maybe once a year and you may be spending five minutes of that year talking to that physician. What can they possibly learn on you outside of, you have something really majorly wrong going on, so you

need help. You don't have this longitudinal tracking over time for what's happening in [00:49:30] your system. You go and you have your blood pressure measured once. You have glucose levels measured once. What the wearable technologies are now enabling for the first time is this longitudinal collection, high density collection of this information that's gonna set a way more appropriate baseline for you to personalize what deviations should the physicians care about in you as opposed to the physician trying to group you into this big population and say, are you an outlier in the population?

If you are it's usually something [00:50:00] very bad is going on. It's allowing us to do your blood pressure five, 10 times a day, check your weight every day. You can wear these watches where it's looking at pulse oximetry, your oxygen levels every day, your heartbeat, your heart rates, your activity levels, your sleep patterns, all of this information that we've been unable to longitudinally assess for cheap before is now possible. Think back to my words on we need really high dimensional data built up around people to make sense of the genome. [00:50:30] That's the data I'm talking about. The fact that these sensors now provide a low cost way to collect lots and lots of information over time, never really been able to be done before will change the face of how we are able to understand the states of an individual, the risk of disease, the subtypes of disease and so on.

Demetri Kofinas: Also the time domain aspect is huge because-

Eric Schadt: Huge.

Demetri Kofinas: That's something that you can't obtain through any type of, even with the greatest sort of, just beautifully matched phenotypic [00:51:00] and genotypic data, a snapshot can't even begin to capture.

Eric Schadt: Exactly right. That is the beauty of this. You'll hear a lot of complaining from others ... A lot of them not wanting this world to come. These devices are like recreational grade. They are research grade, they are not FDA approved. They shouldn't be used regarding clinical care. That may be true if you are looking at sort of an absolute measure, like that one snapshot in the doctor's office. You need that to be nearly a 100% accurate, [00:51:30] so that the physician can interpret it. If you are collecting over time you are now setting a baseline. You are not really interested in the absolute level, you are interested in deviations from the baseline to understand the individual better.

That deviation from a baseline can tolerate a far higher noise profile, where maybe recreation or research grade is fine, because longitudinally over time you are gonna beat down the variation, you are gonna beat down the noise, and that signal will be clear. That hasn't, there are not even many [00:52:00] methods, many statisticians working in biomedical and life sciences who know how to handle that kind of data. I see that as the next big revolution in statistics and how sciences are, how do you even accommodate this longitudinal data in assessing disease risk and so on? I think what we'll find is that it's just a superior way of modeling it.

Demetri Kofinas: There is also a huge privacy component here. I imagine also if there weren't proper regulations put in place dangers and concerns for patients clearly if something is discovered [00:52:30] in your data, what are insurance companies gonna do if they have the extended information? I don't want to minimize that but in the interest of time I do want to ask you, and we can get to that, or we will. I want to ask what, because we are sitting here and we are talking about data and models and all this stuff, and business.

This is really about some aspirational ideas, right? This is about curing some devastating diseases. For anyone that's had to lose someone due to disease, or someone that had to go through something devastating, it's a devastating experience. [00:53:00] It's not small thing. You are potentially changing people's lives, people that work in this industry. I want to ask you that on a very personal level, what do you see, what is the sort of moonshot directly in front of us, what is over the horizon, just over the horizon, and what is sort of out there that you see in the next 20 years or something?

Eric Schadt: I think what is directly in front of us is the ability to now generate for the low cost this really high dimensional data, your genome sequence and so on, and the methods do exist to start [00:53:30] interpreting that information to guide clinical care in some areas. In cancer, areas more primed for this are in cancer. Lots of the more state-of-the-art treatment centers-

Demetri Kofinas: Explain that a little bit, I'm sorry. Why is, let's be clear about that. Why is nothing else more primed for than cancer?

Eric Schadt: Because cancer is a disease of the genome. You get cancer because of these changes in your genome. It's very hard hitting, it's not subtle like Alzheimer's changes and diabetes. It's [00:54:00] fundamentally almost alien. Your cells are now doing something that they should never do. It's very-very easy to see the signal to noise ratio is incredibly high. The super more is low passive next generation sequencing makes it cheap enough, to now generate that on individuals. It's just the ones that's ... The most prime because we can generate the data cheaply, and the insights we can gain from that directly speak to clinical care paths that a doctor would want the patient to take.

Demetri Kofinas: [00:54:30] Let's talk ... In the few minutes we have left, let's talk a little bit about the hurdles and the concerns around privacy and control over data, because that's something that we have experienced in a negative way in our relationship to google, to Facebook, to apple, to all these different companies to TOS, these terms of service. Our data is just out there. Cybercrime has escalated tremendously. The amount of money that's lost to breaches and cyber security is huge. What are those concerns, those legitimate concerns [00:55:00] that people have around data?

Eric Schadt: First I would say the fundamental hurdle again going back to the beginning is just collecting, getting that high enough dimensional data on big enough people for all the modeling to work, because the future will be down the road fully computational models that perfectly simulate the human system.

Demetri Kofinas: You don't ever see the software being a bottleneck in other words?

Eric Schadt: No.

Demetri Kofinas: You think the data is always gonna be behind the software?

Eric Schadt: I think the data will be behind the software and that we are sort of about this scale now. We are managing zettabytes [00:55:30] of data, computing on that scale of data, building and testing these models. That's definitely not a completely solved problem but it's far enough where the biggest bottleneck today is having the right data around the people. Your point on the privacy and the problems with unauthorized release or how companies are gonna use that information, what control does the individual have are all very real, because we need people to want to participate.

I see ... Unlike a google or any other, we are sort of picking on google, [00:56:00] but there are many, yahoo, Amazon, Facebook. They are all the same. You have the digital clickthrough consenting. You are presenting somebody with pages of consenting information that nobody reads. They are allowed to just click through it without thinking about it. We all do that, because we don't want to take time to read the fine print, and we are getting some kind of meaningful service back.

With google, I use Gmail, and it's a good service, because I can send big attachments that store things forever, and it works. I do the clickthrough and I [00:56:30] don't think about it, because I want that service. On the health side, we are not even really allowed to do that from a regulatory standpoint. You have to inform, you have this concept of informed consent.

Demetri Kofinas: Which is meaningful understanding.

Eric Schadt: Right, giving the individual and ensuring they understand what they are actually consenting to. If google had that obligation they would need to be telling you, we actually own the data that you are typing out on your emails. We can use that for whatever purpose we want. Maybe more people [00:57:00] wouldn't be so agreeable to accept those conditions-

Demetri Kofinas: I think it would have evolved in a better way.

Eric Schadt: The right kind of ying and the yang would have been more engaged, right. For us, it's all about just appropriately consenting the individual inform consent, and then enabling them to take control of that information and let them decide, knowing all the risk, like who do you want to share it with, and for what purpose? You just give them that kind of control and then it becomes their responsibility. Of course, we also have a responsibility to protect [00:57:30] that data, to encrypt it, to ensure that it's not easily compromised and the best we can do is using the state-of-the-art industry standards on how you go about that, and having competent IT staff, we've invested very heavily in IT, and big platform data.

Demetri Kofinas: Sorry, I didn't mean to interrupt you. Do you see any promise for blockchain as a technology to help own and disseminate this data and control it as an individual?

Eric Schadt: Yeah, absolutely. I love the idea of blockchain protocols. I think there are few obstacles [00:58:00] that several companies are now trying to overcome, one of them being privacy, with the blockchain protocols everything is sort of exposed and everybody is kind of helping validate, and what does that mean in the context of highly personal health related information. I think there are some hurdles that I know, but I view that as one of the types of technologies that could really help in protecting any individuals that were mentioned.

Demetri Kofinas: Would you say it would be difficult to create anonymity within the distributed ledger, of having to trying to have so many different computer share in that process?

Eric Schadt: Yeah. [00:58:30] That kind of information is see-able by the community that has access to those ledgers and are doing the validations, and how much information is accessible and will inform on an individual, that's still unclear.

Demetri Kofinas: Obviously Eric, you are a brilliant guy. No need to beat around the bush on that. I can assume given the work you do and your general curiosity, you just have some interesting ideas and thoughts around artificial intelligence and where that is going.

Eric Schadt: [00:59:00] Yeah. I'm super mixed on it, because on the one hand I know a lot of what we are doing, if we get to the simulated system that that is going to usher in a new era from designer babies. We will be able to understand what changes do we make to tune these different functional characteristics of the system, their societal implications for that. How do we not create a bigger divide between the haves and have nots? The access [00:59:30] to this technology, the application in reproductive health and stage disease and aging? How do we not let that benefit only the wealthy-

Demetri Kofinas: The fact that people with more-

Eric Schadt: That we are gonna create these super classes that they live forever and have the most amazing kids and then everybody else is a million miles away from that. Having they now serve ... I think there is a potential, unless we as humans become far more purposeful about these changes that we are making, [01:00:00] these technologies we are ushering in and how they come together that we may very well get a world that we don't-

Demetri Kofinas: These are real challenges you are describing. In fact, I'm not sure exactly how one would go about that, because even within the context of an agreement to regulate a regulatory framework, because of the nation state model, you have countries like China that could do something completely different. In fact, there is a concern I'm sure within your industry, which is we want to be competitive with companies that are in other countries with different regulations. There is this tension between what you need in order

to make the [01:00:30] advancements you want or to achieve these ends, and then also the long term fat tails.

Eric Schadt: That's why I think doing these kinds of shows, education, I think we need to get more people thinking, because me as a scientist, the one who wants to drive this hard, because we do want to help people. We are being purposeful on one dimension, but we are not understanding how it's gonna affect these other thousand dimensions. We need more people thinking about that, more people understanding. You need legislators and [01:01:00] your regulatory bodies and academics. It's like how are we gonna transform the state of understanding and democratize it is kind of where I hope a lot of the activity over the next few years kind of takes off to start giving us a shot at getting our head around it.

Demetri Kofinas: I don't think it's an easy question. Many in our audience know I had a brain tumor and I had surgery and radiation from it. Even given that experience, granted ... If someone came to me today and said, "Demetri you have a brain tumor you have six months to live, but if you sign this here all [01:01:30] these changes will happen and you will be cured. Of course, I would take it but in the context of having gone through that and knowing what that's like, even then I'm mixed on this subject, because I am very concerned about it. I have for my reasons touched on some of them, dealing with the incremental effects of this wealth disparity, but there is also the larger question of general AI, which would just usually-

Eric Schadt: Maybe ultimately the ultimate democratizing and maybe that we'll enter digital space, where all [01:02:00] things become digitized. Maybe we are in a simulation now. Maybe that is. I have been somewhat persuaded by that given the amazing progressive in immersive VR. Who knows? There is a reason to be optimistic if that came to be, then all of a sudden everybody is having sort of access to that type of avenue. I think over the next couple of decades it's gonna be a real fascinating thing to watch.

You are gonna see people making direct manipulations, [01:02:30] genome editing of the genome to enhance characteristics that aren't just disease based, because the technology we developed to treat disease can be equally applied to enhance characteristics that people may think are advantageous, better looking, more physically fit. Smarter, better memory. How do we get our hands around that? It's got to be more people thinking about it, there's got to be more legislators thinking, it's got to be educational centers thinking about it. We have to collectively come to a higher level, a higher IQ as a society, [01:03:00] or we are gonna unknown how we'll handle this big divide between the haves and have nots. If that gets bigger it's gonna lead to instability. It's not gonna be a place we want to be.

Demetri Kofinas: Do you ever watch the old star treks?

Eric Schadt: I guess.

Demetri Kofinas: The 60s ones. Remember the Kahn being the super race of people?

Eric Schadt: Yeah.

Demetri Kofinas: From the 90s though. We've already passed that. Dr. Eric Schadt. Thank you so much for coming on.

Eric Schadt: Thank you Demetri.

Demetri Kofinas: I really appreciate it.

Eric Schadt: My pleasure.

Demetri Kofinas: That was my episode with Dr. Eric Schadt. [01:03:30] I want to thank Dr. Schadt for being on my program. Today's episode was produced by me and edited by Stylianos Nicolaou. For more episodes you can check out our website at HiddenForces.io. Join the conversation at Facebook, Twitter and Instagram at @hiddenforcespod, or send me an email. Thanks for listening. We'll see you next week.