JAMES ALTUCHER'S

INVESTMENT NETWORK

Monthly Mastermind:

Igor Tulchinsky & Chris Mason Official Transcript

JAMES ALTUCHER: I think the biggest job of the next generation, this is what I would do if I was figuring out what to study and train in is predictor. We've seen examples of this. We saw the movie Moneyball where they used data about baseball players to predict which baseball teams would do the best. And I used to write algorithms for the stock market to use data to try to predict how stocks would act, and we're going to talk about that more today. And people use data also now to analyze the genome to predict what diseases someone might have and on and on.

But the technology has gotten so good, so fast. I'm happy to talk with the two authors of the book The Age of Prediction, Christopher Mason and Igor Tulchinsky. Now, Igor is interesting from the financial perspective. He has a $7 billion hedge fund which analyzes millions of pieces of data around the world to predict stocks. To predict simply what's going to happen with stocks tomorrow, or an hour from now, or 10 seconds from now?

And Christopher Mason uses data from the human genome to study what diseases we can start curing, what sicknesses someone might have, or what traits someone might grow up with. So, I wanted to know what is the state of this industry. How much can we really predict? How can we get better at it? What are the limitations? And then we just had a fun time while I pitched different ideas. So, here's Igor and Chris, authors of The Age of Prediction.

JAMES ALTUCHER: Chris, I'm amazed at all the things you were able to discover about people by analyzing their genetics.

CHRIS MASON: Yeah, basically, it's a predictive algorithm that's in every cell. So, you have bits of DNA, RNA, proteins, and you leave these everywhere you go. So, I think at the forensics chapter you're probably thinking about maybe, or even just the cancer diagnostics we can do, DNA is extraordinary.

JAMES ALTUCHER: So, the book is called The Age of Prediction: Algorithms, AI, and the Shifting Shadows of Risk is the subtitle. If I can try to summarize at a 20,000-foot level, it seems like there's three types of prediction. One is where you use statistics/AI to model things that could either be very predictable or somewhat predictable like things ranging from insurance risk to cancer risk using genes to stock market predictions.

Then there's the kind of prediction where it's just someone's opinion like Rifkin analyzing what the economy's going to be like five years from now. And then there's the kind of prediction where it's pretty definite. So, the solar system's going to eventually collapse. I can predict that, and I know 100% chance I'm correct. Would you say that's roughly three categories of prediction?

IGOR TULCHINSKY: Yeah, based on facts, based on optimism, and based on reality.

JAMES ALTUCHER: I like that. What about based on pessimism?

IGOR TULCHINSKY: It's the same as optimism, right? Because you just flip the side.

CHRIS MASON: It just goes negative direction, but you can still use pessimism and look at the same facts, and it could almost be depressing to you if you think, "Oh, we're doomed because the sun will engulf the earth in a few billion years." But you could think, "Well, no, that means we know when we got to get moving by." It could be exciting and get you moving.

JAMES ALTUCHER: Exactly. So, you point out the exponential growth in data. Maybe you could describe that a little bit like how much more data we are generating now than even 10 years ago, and what that means for prediction.

CHRIS MASON: I'll jump in. First is I think the amount of data certainly in genomics and just the ability in biomedicine to generate data is what's often been called, at least in genetics, is genomical amounts of data. Most people think of astronomical amounts of data as being really big and involving exabytes or yottabytes of data, which is trillions of terabytes.

But it's actually genetic data and genomic data are now eclipsing the amount of data made by telescopes and astronomy. So, there's a paper that just described this called Is It Genomical Data or Astronomical Data? Which One's Bigger? And concluded that there's actually more genomic and biomedical imaging data than there is astronomical data.

So, when you think of really large, you have to think of things in trillions of terabytes. Not quite today, but in the near future. And that really basically it means every day that you wake up, there's more data than any other day in human history.

JAMES ALTUCHER: Let's do a little thought experiment. So, let's say I wake up and now I can sample my DNA in seconds and do some diagnostics right away, or some AI or some statistics can do some diagnostic right away. What theoretically could I learn about myself this morning from my DNA?

CHRIS MASON: If you grab your DNA, so there's a lot of DNA in your body. About half of it is actually microbial DNA that moves and changes and evolves, and sometimes every 20 minutes the bacteria are dividing. So, you'll learn about any changes in your microbiome, which are the small creatures in and on and around you that have moved. So, maybe you could pick up obviously a pathogen like flu or COVID, SARS-CoV-2 which causes COVID. You could of course get sick.

But a lot of them are actually things that are the anchor for a full ecosystem that itself is a little pharmacy. So, you could see if there's any changes in your gut microbiome if you have problems with your gut, for example. You can look at your epigenetic changes, where it's not your DNA but also how it's packaged and regulated.

And of course, you can get new mutations. Every day you get mutations and most of them are harmless, but some of them could be the beginnings of a cancer that you could see if you saw it that very first day.

JAMES ALTUCHER: What about from the DNA itself? And this is not something you would find in the morning. When you sequence my genome or whatever, you would get all the hard-coded things about my DNA. What can you see from that?

CHRIS MASON: Well, from there you can tell a lot about you like your ancestry, for example, do you have any Jewish ancestry? We had a fun part of the book where we talked about Igor as Jewish, but then at the end of the chapter, he got even more Jewish because the databases got updated. So, we can see the databases… And this is a good lesson of prediction is that your predictions are only as good as your training data and your databases. And when you change the databases, you'll get usually improved predictions, but sometimes it can go the other way.

So, we can look at ancestry, your risk-taking likelihood, how fast you process caffeine or other drugs. So, you can really be predictive about how and what way drugs and molecules will be processed in the body of any person.

JAMES ALTUCHER: So, some of this seems predictive, some of this you know for a fact. For instance, with the DNA, there are some genes where... And I'm going to simplify it with my language, but if they're on, you have Tay-Sachs disease, for instance, and if they're off, you don't. So, it seems like some things they had enough data that they were able to figure out which single mutation genes cause which diseases.

And some data though is more predictive like, oh, is this person more likely to be happy or sad, or Jewish or not Jewish? And you do that by matching tens of thousands of humans who have sequenced their genome. You know what they... This person was Jewish and happy, and this person was something else, and also not happy maybe. And then you can start to build together probabilities based on a new genome.

CHRIS MASON: Yeah, in a nutshell, that's in the ballpark right, and basically, you look for differences in the phenotype or what people express as a trait, and you compare that to the genome. And it could be everything from height, for example, which is there's no one gene for height, there's not even two. There's probably several hundred genes that really mediate how tall you are, but it is very heritable. If you look at tall parents, they'll have tall kids, and the inverse.

So, it is very heritable but very complex. It's what's called polygenic, meaning there's more than one gene that influences that trait in a highly heritable way. And so, as we found more of these genes, they'd get built into the models and we can predict to within about an inch or so how tall you'll be. So, if you take a baby at birth, sequence the DNA, we can get down to within an inch of how tall they'll likely be.

JAMES ALTUCHER: And then how far are we from the technology to manipulate a gene at birth or a sequence of genes at birth to change someone's height?

CHRIS MASON: We're actually doing it, not for height, but we're doing it for disease and modifying DNA as we speak. So, you can actually modify... Embryos have had disease genes removed, for example, for hypertrophic cardiomyopathy or heart disease gene. Or even in an adult, there's been treatments to get rid of beta thalassemia by doing gene editing in the person's body as an adult for sickle cell and beta thalassemia, these blood disorders.

JAMES ALTUCHER: Now, are these single-gene disease?

CHRIS MASON: Correct, single gene.

JAMES ALTUCHER: So, with multiple-gene disease, this is where the data is immense. The possible permutations of which genes could be... Let's say height is caused by a hundred different genes out of what, 32,000 genes or some outrageous number?

CHRIS MASON: About 60,000 total, yeah, yeah, yes.

JAMES ALTUCHER: So, the permutations are in the... I don't know, quintillions, quadrillions. So, it's impossible to use statistics or a computer for that. Is this something AI could start to figure out when there's multiple genes involved like feeding it through neural networks the way they did with ChatGPT?

CHRIS MASON: You could basically feed the data from basically millions and millions of patients and their clinical metadata and their traits, and essentially learn what are the new signatures that are driving some of these changes. But it wouldn't just have to be genetic data. It could also be what was in your data or what was other factors in the environment. What else mediates that last few percent you can build into these models as well?

JAMES ALTUCHER: But how far are we from really the technology to really understand things like height or intelligence or cancer? All of these factors that involve hundreds, maybe thousands of genes.

CHRIS MASON: Well, in some cases we're very close I'd say. Height, for example, is pretty well teased out. Even autism risk, even though autism is complicated, it's several hundred genes, many of them are now consistently being identified, so we can explain a lot more autism than we could certainly 10 years ago, and almost couldn't barely do it at all 30 years ago.

So, I think even complex diseases or complex traits we can now explain pretty well, and you can edit them. You can edit dozens of genes at one time, what's called multiplex editing. In pigs, for example, they've done up to 60 genes at a time. There's new trials that can give you up to 100 edits all at the same time all over the genome.

The worry though is it's not perfect, so it's like sending someone with a bunch of erasers through your book, and if it's correct, you would very precisely do changes in the text of life. But if it's messy, then of course, it'd be hard to read the book because you've made too many mutations. So, that's what we're working on now.

JAMES ALTUCHER: Because the more genes involved, the more you could have side effects in editing them. Like what if some set of the genes for height are also related to genes for, I don't know, some disease or whatever, cancer?

CHRIS MASON: That is the risk. There are new methods for CRISPR, the ones called prime editing where instead of breaking both strands of DNA and swapping out a chunk, and then having that occur sometimes off-target effects, meaning not where you want them to be.

Prime editing breaks only one strand and actually is much more precise. So, you could use some of those methods, or there's actually a quest by many companies right now and a lot of money being invested to try and find newer and more precise methods for editing. But the technologies are here today and they're only going to get better.

JAMES ALTUCHER: So, Igor, this strikes me as... You've been involved in, for instance, predictive algorithms for stock market predictions. This strikes me as a little different than that because the difference between the genome and its relationship to diseases and traits in the body, that's probably accurate. Once they figure it out, they know. Once we know these genes affect height, we know forever that those genes affect height.

But with the stock market, the more people know something, the less likely it is to work next year. For instance, if you're going to predict new additions to next year's Russell 2000 so you start buying the stocks now, well, once everybody starts predicting that, it's too late to use this algorithm.

IGOR TULCHINSKY: Right, that's right. In genomics, what you figure out does not affect the subject, but in finance, the fact that you figured it out is going to change the way the system behaves eventually.

JAMES ALTUCHER: I mean I take some blame of some algorithms that I used to use stopped working after I wrote about them because I have the misfortune of loving trading, but also writing. So, I had an algorithm that it might have been... It seemed statistically significant to me. The past 80 times roughly, the cues gapped up between 0.4 and 0.6%. You could short it at the open, and they would be flat at some point within the next hour or so. And it was like an ATM machine for me. Every time it happened, I made money until I wrote about it and then it was actually random after that.

IGOR TULCHINSKY: Well, you made the market more efficient on the other hand.

JAMES ALTUCHER: I'm a hero.

IGOR TULCHINSKY: You're a hero.

JAMES ALTUCHER: So, there's this spectrum between, okay, given data we can know some facts, for instance, on the genomic side. Or given data, we can predict, but we don't want to tell people our predictions if we're making use of it. So, there's a spectrum there.

So, like Moneyball, the book by Michael Lewis predicting baseball outcomes, that's more like a stock market-style prediction. Because if you could draft people who are good at walking then everyone's going to draft it and the arbitrage goes away for teams.

IGOR TULCHINSKY: That sounds right. That sounds right.

JAMES ALTUCHER: What are other categories like that?

IGOR TULCHINSKY: Maybe even traffic. If Google algorithm was routing everyone in one direction, soon it will not be empty, and where there is no traffic, there will be traffic. And possibly most algorithms that make predictions in systems in which the user is participating will be like that. It's only if you are predicting something away from your ability to influence it that it should not change too much.

JAMES ALTUCHER: Did you ever get nervous in your hedge fund career that you just were going to run out of algorithms? That eventually all the... Because now there's 30,000 PhDs in every hedge fund trying to find these discrepancies in the data or in the outputs. Did you ever worry all the arbitrages would be gone, the market would be smoothed out, and that's it?

IGOR TULCHINSKY: In the beginning, I used to worry about it, and then I would have these interviews, and everybody was worried about it that the market would become efficient. But it never has, and actually, simply logically thinking, somebody has to make it efficient. So, somebody's going to be there making it efficient no matter what. It may just not be you, that's the problem.

JAMES ALTUCHER: But can't my predictions also be mean reverting? So, if a prediction worked regularly for a while because presumably, it modeled some mob psychology. And then if it stops working for a while, won't it mean revert and eventually work again?

IGOR TULCHINSKY: Yeah, and when it does, it'll come back to life. It'll get turned back on.

JAMES ALTUCHER: Okay. Yeah, that makes sense. And also, obviously, you talk in the book a lot about insurance and insurance risk. And the entire insurance industry is built on predictive modeling, but that's been the case for let's say hundreds of years. What's new in predicting human behavior now that has helped the insurance industry?

IGOR TULCHINSKY: It's the immediacy and the real-time data. You can use real-time information coming, let's say from people's driving to adjust their rates. You can use body sensors to change outlook on the health of an individual and so on. I imagine it used to be that you would fill out a form once a year and that's where it stood. Now, there's more and more data coming in, so the insurers are getting a clearer and clearer picture, which theoretically, is good for everyone.

CHRIS MASON: Car insurance, for example, they put a little device that goes in your glove box, basically. They say, okay, it tracks your speed and your GPS coordinates and looks to see if you're speeding basically. But when we got a recent insurance, my wife didn't want it in the car. She was like, "I don't want that thing in our car." So, we purposely took the higher rate of insurance just because she didn't want that in the car.

So, the insurance companies now are saying, "Oh, you get a discount if we put this device in your car," so they can build better models of you. But you could always just not take it, I guess, but then they make you pay more for it.

JAMES ALTUCHER: Yeah, have you ever read the book 2041 by Kai-Fu Lee? So, Kai-Fu Lee is a big AI technologist from way back, essentially the father of speech recognition, and he wrote this book, basically a bunch of scenarios about how predictive AI might work in 2041.

The first story he uses insurance as an example where this family gives over all of their emails in exchange for a discount, but then the insurance company could model them better. Their insurance rates spiked because, I guess one member of the family from her emails, it could be determined that she was in a, let's say not pleasant relationship. And so, her risks of accidents they knew would go up according to their data. So, it was pros and cons to the enormous amount of data we can now capture to use for prediction.

But my question here is this actually is closer to the genomics model where just because you know something it still might not prevent it from happening as opposed to the stock market predictions.

CHRIS MASON: You're not necessarily fated in the sense that most genetic risks are probabilities, but some things are really hard to avoid like Tay-Sachs disease or cystic fibrosis where you'll almost certainly get some... Or Huntington's disease. There are some diseases where it's going to be really hard to avoid.

But in the book, because of the CRISPR and different genome modification systems, you're no longer subjugated to the shuffle of genetic lottery you got as an embryo. You can in theory modify it or tweak it or think about what you hand down to the next generation. For the first time really in ever, we have the ability to tweak what is that risk.

JAMES ALTUCHER: I feel like it's still... Yes, again, for the single-gene mutations that are causing diseases, you can turn an on-and-off switch and get rid of the disease using technologies like CRISPR. But most things are more complicated, and I'm just wondering, what are the first complicated things that we're going to be able to solve?

CHRIS MASON: I think-

JAMES ALTUCHER: Like heart disease.

CHRIS MASON: Yeah, heart disease would be one, and there's mutations that can drive this, and there's even known genes for hypercholesterolemia that we could target. Some of those have already been targeted actually, in terms of... Again, it's usually one or two genes that are being targeted and there's more than one.

So, in those cases, or for example, if you look at certain kinds of cancer, they're driven by... BRCA1 and 2, for example, a lot of breast and ovarian cancers are driven by a handful of genes. Probably the top 30 genes alone would explain 90% of the cases or so, 95% of the cases even for ovarian cancer.

So, if we know what those genes are, you can constantly be scanning in the blood for them to see do you see a spike of any mutation and say, "Aha, I see it." And we're going like whack-a-mole, take it and get rid of that mutation. And if another mutation comes in a different gene, see it, and then go after that target.

So, I think it would end up being... You rarely would need to go after, say 15 genes at once. You'd probably do it over time for more complex diseases like cancer. But for some things like height, if you think you're going to be, for example, really 4'2" and you wanted your kid to be taller. So, again, this sounds hypothetical, but something like this will probably happen where someone says, "Okay, I have a safe way to make sure your kid is tall, and do it in an IVF clinic."

And it hasn't happened yet, but the closest thing is a company called Orchid which is doing this for embryos. They sequence the genome of each embryo and then you pick the one that you want based on that selection.

JAMES ALTUCHER: Because you could predict, okay, this embryo is going to be a female, tall, athletic, high IQ, and this embryo... And let's say with odds on each one, but pretty good odds. And this embryo there's odds that, oh, low IQ, not athletic, male. So, I'm going to abort that embryo and give birth to this one.

CHRIS MASON: Yeah, which is happening today with Orchid at least. There's other ones that are also coming on into the market that are trying to guide IVF basically, but it's embryo selection, and that's just-

JAMES ALTUCHER: I see, so embryo selection before embryo modification. It's a little easier to do that because modification we don't yet... We're just making guesses on the odds. We don't know 100% chance this person's going to be 6'3" but it's like a 60% chance, which is better than these other embryos. So, then we don't have to take our risks with modification. What do you think in China they're doing?

CHRIS MASON: What they're doing there is a lot more of the somatic methods I've seen where they're doing things in your body as an adult, modifying your cells. They've been looking a bit more at embryos, and also, actually being much more aggressive with which new modified cell therapies. Basically, genetically modify a cell, reinfuse them into the patient, and then have a targeted therapy that happens.

But it's targeting something that, from our own publications, we wouldn't recommend. Because if you target what you think is only on a cancer cell, but it turns out it's also on 20% of all the other cells in your body, that'll probably be very painful because you suddenly unleash all these angry immune cells that are attacking what it thinks is on a cancer but is also on regular cells.

JAMES ALTUCHER: Yeah, it seems like this could be really powerful for anti-aging depending on which model of anti-aging you believe. If you attack telomeres and change the genes in them so they don't shrink over time, or you are able to inject new ones in and they attach to... I don't know how it all works, so they attach to your cells or whatever, it seems like this could be really great.

These things called Yamanaka factors that they're researching in Asia, I guess, but there's overlap with cancer. Apparently, the more you get this kind of treatment, the more likely you are for cancer, so there's all these risks.

CHRIS MASON: Yeah, absolutely. And so, I think you have to balance what is going to be the likely benefit from it from any possible side effects. Do no harm is the basis for most medicine, and in some cases, we might not know though. We think we know, or we looked in a mouse, but we don't know yet in a human. So, that's why clinical trials always start small. You start with 10 people, maybe eight people. Start small for people that really need it and then you slowly expand.

JAMES ALTUCHER: What about the kind of stuff that Palantir does? So, given a set of bank transactions and a bank customer, they can make a guess or a prediction as to whether this bank customer is a terrorist or not, for instance, or is involved in some kind of financial fraud. And again, this is more related to the stock market prediction stuff because, as you know more how they're predicting you, you could modify your behavior. But how close are we to really modeling human behavior like that?

CHRIS MASON: I mean credit card companies do that today, I think. They'll look at any spending patterns, look at any changes in behavior, try and guess whether... Well, one, just to guess whether it's you. Did someone take your credit card, or did someone grab your phone? And then the other thing that they'll do is just look to see are you a risk in some other capacity.

I mean the other thing it does, for example, is iTunes or Spotify try and make a playlist based on what you've listened to before, which is great. But then when my daughter got my phone and started listening to all of her songs, it totally screwed up my algorithm. So, now it's no longer my ideal playlist obviously.

JAMES ALTUCHER: So, given that that kind of prediction like the Netflix or Amazon-style like, "Since you bought this, you might like this," that's been around for a while. Yes, I'm sure it's improved, but what's really cutting edge with that? I get it that with genomics this mapping all the permutations of possible data to real diseases and human traits, that's important. Obviously, the stock market is an immense problem that can never be fully solved. What things are blowing you away on the social modeling given the sheer amount of data we now have as compared to 10 years ago?

IGOR TULCHINSKY: So, the movies that get recommended to me I actually like a fair percent of the time. Life's getting easier, less thinking to do.

JAMES ALTUCHER: Yeah, and also, I guess I used to get called all the time by credit card companies saying, "Oh, this is suspicious behavior," but it was just me being me. Now I don't get as many calls. It seems like they are better at modeling if a credit card is you or not.

CHRIS MASON: You might still be a suspicious person just by your weird habits, yes, but at least it knows that now and is-

JAMES ALTUCHER: Yeah.

CHRIS MASON: Yeah, yeah.

JAMES ALTUCHER: But okay, what do you do in situations like March 2020 when the advent of COVID happens, and you talk about this in your book? The market falls, it's the equivalent of what Nassim Taleb would call a black swan event. The market falls eight standard deviations more than normal in a short amount of time. Something that should happen one in a trillion times, and yet it happened.

What do you do when there's really no model? And again, the stock market's considered a fat-tailed curve as opposed to a bell curve. How do you take into account situations like that where you can suffer significant loss treating it like a bell curve?

IGOR TULCHINSKY: You divide things into ripples and the waves. So, the waves are things like you described that are these gigantic events that are specific, and you can see them, identify them. And the ripples are the stock of a toothpaste company moves .1% when something else happens. So, we create the ripples, but we stay neutral to the waves.

When an event like 2020 happens, it shakes us somewhat, but we're more or less neutral, so we stand through it. But where there are waves, there are ripples too, and the ripples start going in all kinds of different directions.

JAMES ALTUCHER: I see, that totally makes sense. For instance, if something big is happening that's affecting the entire market over a long period of time, whether it's a day, or weeks, or months, that's not your business. But let's say for 20 seconds the Canada markets deviate from the US markets by a wider spread than usual, you could say, "Okay, within the next 20 seconds they usually snap back." And you could play things like that regardless of the larger wave that's happening in the markets.

IGOR TULCHINSKY: Yes, it's like that but it's not only the time. The ripple can have a long duration, but it's just very weak so that nobody else is really trading but you.

JAMES ALTUCHER: I see, so in AI, of course, there's this difference between unsupervised learning and supervised learning. Much like how ChatGPT was built initially, you could use unsupervised learning to find where the AI finds that the context of some patterns of language is related to some movements of stock. We don't know what the connection is, but there seems to be a connection. And then from there, you could just start figuring it out.

IGOR TULCHINSKY: You just need the statistical relationship, you don't need to understand why it works. It may not be possible to understand why it works, and by the time you understand why it works, it won't work.

JAMES ALTUCHER: Right.

CHRIS MASON: Yeah, I mean all the things you do with the data, I'm trying to think of other layers of that data. So, for example, where do we see higher rates of different diseases, or cancer, or infections could be related to what is your... Are there toxins nearby? A geospatially informed view of healthcare could help you stratify risk, and you can find causes and essentially why some people might be getting sick.

This is something you can maybe use all the Google imaging data, for example, from Google Earth, and look for trends there. A simple thing as the number of trees or night lights in a neighborhood can influence health in some ways. And so, you could use that. I don't know if you could make a lot of money on that, but you could at least stratify risk and help people stratify their outcomes.

JAMES ALTUCHER: I wonder if there are things that we could look at that we're not looking at that could help us decrease let's say a spike in accidents in some geography or things that we just haven't even thought of because we don't know the connection... We don't really comprehend why there would be a connection but there is one.

CHRIS MASON: Yeah, and that's the really good thing about a lot of the tools for stratification. All these AI tools will find the patterns and then you can build that into a model. You can essentially leverage that to make a better prediction. We actually do this, for example, when you sequence a potential pathogen from a sample, from a urine sample, for example. We look at all the facets of the data, not just what species is there, but statistics on the fragments of DNA that came out or the pH of the urine or other factors that could better diagnose a UTI.

And everything goes into the models and essentially, we don't even need to know why the model gets better, but if it can predict better what pathogen is present, then we can use it. And actually, some of these are under review now by the FDA to really embrace some of the AI algorithms because they work really well, and they'll lead you to a better way to do diagnostics and care.

JAMES ALTUCHER: What about just pundit predictions? So, somebody goes on CNBC and says, "Well, I think gold is going to go up because of geopolitical stress, blah, blah, blah." Do you think humans have gotten better? Given more understanding of history, more data about recent events, more opportunities to predict and see how those predictions turn out. Do you think humans have gotten better at being essentially pundits?

IGOR TULCHINSKY: The answer is paradoxically no because the more the ability to predict things improves, the more people lean on those predictions, the more they're used, and what remains is a more and more unpredictable world that gets harder and harder to predict. So, by the time somebody is saying something about gold on CNBC, everything he knows has already been figured out, and other more notorious pundits on there that there's probably nothing left to pundit about.

JAMES ALTUCHER: Well, what about a macro trader like someone like George Soros who famously predicted the collapse of the pound in the early '90s? I think it was 1991 or 1992, and was it just luck that some macro traders succeeded, and others didn't, or was there something else? Did they have some special insight that maybe has disappeared from the markets now?

IGOR TULCHINSKY: I think they had insight and understanding and there was not too much competition for high level of insight, but these days there is.

JAMES ALTUCHER: Yeah, it seems like that's the most... And the trading arms of all the banks seem to be very quant-focused because they do all the high-frequency trading and so on. So, again, definitely in medical there's opportunity. You have a lot of data, so now we can start figuring out which multiple genes relate to what characteristics and traits. Although there it seems like a math problem. You have to deal with these exponential-size math problems that computing can't do.

Is quantum computing a solution for figuring some of these things out if and when ever there is quantum computing? I'm not sure there ever will be, I'm not sure I understand it, but is that a solution to the exponential problem?

CHRIS MASON: If it does what it's supposed to do, it could help and give us just that much more compute capacity on the planet. And so, that would certainly help but I think a lot of it'll also be some of the testing will be done on the ground, or you could do some of it with model systems. But I mean I would be the first one to jump in line if we had solid quantum computing up and running. It'd be great.

IGOR TULCHINSKY: There are classes of problems that quantum computing can crack and classes of problems that still remain unattainable.

JAMES ALTUCHER: Really? What's a type of problem that a quantum computer can't crack?

IGOR TULCHINSKY: These days they're selling encryption algorithms which are quantum-proof, so that's one type of algorithm.

JAMES ALTUCHER: Which is a good example because classic encryption right now, let's say the way Bitcoin's encrypted, that can be solved by a... Factoring a 100-digit prime number cannot be solved by a thousand supercomputers linked together, but a quantum computer can do it in a second. So, that's a classic example, and now you're saying there's algorithms that could make cryptography quantum proof.

IGOR TULCHINSKY: There are, there are. There's no computer that can solve all problems.

JAMES ALTUCHER: Because what I worry about is are we hitting a point of let's call it peak data where we have the maximum amount of data for certain categories that we can basically handle because the computers are not fast enough. And they're not fast enough not because the chips are slow, but because mathematically, it's too exponential a problem.

IGOR TULCHINSKY: Yeah, that can happen. The data, the rate of the growth in the data may simply exceed the computing power's ability to digest the extra data.

CHRIS MASON: And I don't think we're there yet, but we certainly could because we're generating so much data. Yeah, maybe, but peak data indicates that the data will have less utility or that it will have peaked past the ability to use the data. So, I don't know if I'd call it peak data, it's just unwieldy. We've reached past being able to wield and manipulate data efficiently to only having various degrees of efficiency. Because I think the data, assuming it's clean data, should still get more useful as you get more of it and over time.

JAMES ALTUCHER: Right, so you'd have to come up with more categories that you're studying in the data. But again, genomic data seems like it's there. For 15 years we could predict Tay-Sachs and other single-gene mutations like what single genes cause which diseases.

But I feel like for something like possibility later in life of stroke or predicting IQ. Something like that which requires hundreds of genes, we're never really going to be able to... The data's there and it's possible. The algorithms are there, but our computers are never going to be fast enough to solve them.

CHRIS MASON: I mean maybe today, but I wouldn't say never because I never say never because I think there still could be... In 50 years, there could be something that's even beyond quantum or that's some other variation of a computing. Or even just efficiencies of the algorithms could be improved in ways we can't imagine now.

So, maybe in the short term, yeah, but I think long term I would say no. I mean if you go back 200 years ago, it would have been inconceivable that people would routinely fly through the air in airplanes. No one would have believed you and no one believed even the Wright brothers for a while. So, I think a century is a long time with current humanity.

JAMES ALTUCHER: Well, it's really interesting you say that about the algorithms might improve. That's an area I haven't thought. Obviously, over the years, over the centuries, statistics has improved. So, instead of just trying to match something against a normal curve, now there's all these very sophisticated algorithms for speech recognition, vision recognition, and so on. How much more do you think the math can improve because that also seems to be very... We might be at peak math in terms of how much math is actually useful that's coming out of academia.

CHRIS MASON: Yeah, like math departments. There's no new... I mean there's a lot of work on essentially reconciling quantum mechanics with Newtonian physics, and that is still something that should be solved at some point. But there's no new kinds of numbers being discovered or entirely new kinds of calculus. Most of that's been done since the 16th century.

And so, I think there are some... I mean Newton probably would have argued he was peak math. I don't know, maybe a long time ago with calculus, but-

JAMES ALTUCHER: But even statistics has with, again, the rise of the need for pattern recognition, I feel statistics has evolved in the past 30 or 40 years. These hidden Markov processes and other techniques to really, through sophisticated pattern matching, and I just wonder can that be improved?

IGOR TULCHINSKY: Yeah, depends what you call an algorithm, right? Maybe the basic algorithms can't be improved and don't need to be. But if you look at something like AlphaZero or ChatGPT as an algorithm then yes, from time to time a groundbreaking algorithm does appear.

JAMES ALTUCHER: Yeah, but I wonder how much of that was, okay, more advances in adverse neural networks versus the speed of computers finally... Look, the speed of computers were there to solve more or less computer vision 20 years ago. But I feel only in the past 5-10 years it was fast enough to handle large language models like ChatGPT, and that was purely a speed thing. How much was really an algorithm thing?

IGOR TULCHINSKY: I think there were elements of both.

CHRIS MASON: I agree.

JAMES ALTUCHER: Yeah, that makes sense.

CHRIS MASON: And also, just the amount of data. You couldn't build large language models until you had lots of data to look at. So, all those things together coming, the data, the algorithms, the compute, then you could do it. But you really couldn't do it 15 years ago.

JAMES ALTUCHER: Yeah, and actually, to your point, we didn't even have the data. We didn't have all written text up until last year stored anywhere in one easy-to-use place. But even that, with the speed it took a bunch of supercomputers a year and a half to crunch the large language model that's now ChatGPT. And then another year and a half of supervised learning. So, it'll be interesting to see how that speeds up.

So, Chris, I know you're interested mostly in predicting breakthroughs in medical technology, and Igor, in financial and stock market predictions. What other things would you like to predict?

IGOR TULCHINSKY: I think the methods you can use them anywhere. What we're doing in ours, we're actually predicting data itself. Data as this property that you mentioned before that when you predict it, the data doesn't change from the fact that you predict it. So, we're predicting data, and you can remodel most things as data and then you just start going into different industries and how much can each industry be improved through prediction, through algorithms, and through just getting rid of that 90% of the work that's mechanical in nature if you consider that ChatGPT is a mechanical thing in the end.

JAMES ALTUCHER: Do you think I can predict what a hit song will be? Let's say I take every hit song of the past 20 years and feed it into my statistics/AI machine. And then I use AI to create a video with a beautiful woman or guy, and boom, do I have a hit song? Do you think that's possible?

CHRIS MASON: Good beat, good voice. I mean there are some things that have catchiness to it. There's definitely signatures of songs that are poppy if you will, and they're popular. So, I think you could build the model and maybe you could get 99% true, but I don't know if you could ever say 100% any model.

IGOR TULCHINSKY: You may have to model influencers and feed the song to the right influencers to get it popularized.

JAMES ALTUCHER: Yeah, that's a good point. But maybe I can create an influencer using modeling. So, okay, here's everything said on Instagram over the past year that got a million likes or more, and break down those. And then here's what every influencer looked like. So, now I'm going to come up with the average super influencer, and then boom, now I'm going to feed that influencer a song and make a record label and use that.

IGOR TULCHINSKY: True vertical integration.

JAMES ALTUCHER: Yeah. People wonder if this rise in predictability is going to bring down creativity, but I think it's going to bring up creativity because now it's going to free up your resources in some ways to come up with even more creative ideas. What do you guys think?

IGOR TULCHINSKY: It's certainly easier to ask questions and get very quick answers, which frees up creativity because it saves time.

CHRIS MASON: Well, it's like any new technology, it can be a tool or a weapon. In this case, it could be a great tool for creativity. AI, of course, people are afraid of AI because it could be in theory weaponized, but on the creativity side, I think it's a time-saver, it's going to be inspiring. You can craft something probably then with Stable Diffusion or DALL-E other tools. You can just describe the landscape you're imagining, and it creates it for you, and you can build from there very quickly.

And these amazing portraits. I think it's phenomenal. It lets you do what I always wished I could do as a kid was describe something, a scene, an idea of say a rabbit with 12 different tentacles that was playing the harpsichord and also juggling pool cues. I could never really draw that, but I can get an AI algorithm to make that in five seconds, right? So, it's phenomenal.

JAMES ALTUCHER: But maybe there's a problem there where... One of the reasons why people always say, "Oh, the book was so much better than the movie," is because when you're reading the book, you're constructing the movie in your head instead of it being... And then suddenly you see the movie and you're so disappointed because it wasn't good as that movie you built in your head. But now we're going to be able to see basically the movies that we build in our head much faster.

CHRIS MASON: Yeah, I think which is still... It could be good and bad, but I think mostly good because you're getting the movies out... You'll see what's present faster, but then also you could have 55 variations of it. You can change the seed kernel for most of the AI art, for example. So, you make 50 versions of the thing you were just thinking, and you actually might then imagine things that you weren't quite originally thinking.

So, you don't just get one imagination, you get 50 of them, or you use some pruning, maybe prune it down to 25 that you like. But I think that's... Just imagine if you could have 50 brains instead of one, you can have that today, which is pretty amazing.

JAMES ALTUCHER: So, what do you think? We're in the age of prediction. Obviously, I should tell all my kids to be data analysts because that's going to be just this huge profession for the next 20... I'm making that prediction. That's going to be a huge profession for the next 20 or 30 years. 20-30 years from now, what do you predict we'll be seeing in our predictive abilities and how we use it in society that will just blow our minds? Curing cancer by the way won't blow my mind because I expect that.

CHRIS MASON: Yeah, that's supposed to happen. Yeah, it's got to be... But it could be every toilet will be monitoring you. Every morning you get a little update report on every molecule in your body. You would get information about the environment around you, around your home or apartment.

I think we'll see prediction coming even from other planets. For example, the Perseverance rover landed on Mars. It took too long for a signal from Mars, so it had to use image recognition software during a landing to get there. So, we've started to see prediction algorithms and tools even send data back from other planets like Mars.

IGOR TULCHINSKY: We're able to get news from the future because it'll be mostly predicted.

JAMES ALTUCHER: On a micro scale, I can say, "Oh, this person crossing the street moves like someone who's going to rob a car in the next day. So, Minority Report-style predictions, the movie with Tom Cruise. What types of news events do you think will be predictable?

CHRIS MASON: Maybe elections. Elections might... I mean we saw this with Cambridge Analytica, they might be tweaked before they even happen. So, you might know the future because you've made the future. But it's possible.

JAMES ALTUCHER: But like-

IGOR TULCHINSKY: I think if you examine news headlines and just examine the news, you will find patterns already that so much news follows other kinds of news and is predictable and so on and so forth. But nobody is really putting that into a news service these days.

JAMES ALTUCHER: That's fascinating. So, what you're doing there is you're disconnecting news a little bit from reality, which is what newspapers probably do anyway. And you're saying tomorrow's headline is more based on today's headline than in the actual events that happened today.

IGOR TULCHINSKY: Yes, yes. So, you may be able to print tomorrow's newspaper better than the actual paper what's going to be printed tomorrow because tomorrow's paper will have some noise in it from new sources, et cetera, et cetera. But your prediction is based on mathematics.

JAMES ALTUCHER: Well, The Age of Prediction, such a fascinating book. I really do think the job people should be preparing for is data analyst because that's going to be used in every single industry. More than prompt engineers or AI coders because AI is going to write its own code. Being able to understand what data to look at and why and how to make use of it. Whether it's the medical industry or sports or stocks or insurance or art, this is going to be such a valuable skill to have, and it's a just beginning field. The creativity there is going to be amazing.

But The Age of Prediction is like a guidebook to what's happened and what's going to be happening and all the ways people use prediction technology, and such a great book. How did you guys team up to write it? Why did you write it together? How do you know each other?

CHRIS MASON: We originally met actually at a lunch at Cornell's campus and then just started brainstorming about data. And then started walking through the lab chatting about the use of data for medicine and overlapping with finance, and it just became exciting to think about more ways to do partnerships, brainstorming. We have a fellows program that goes between WorldQuant and Cornell as well, so people can go back and forth. They're kind of this nice exchange of ideas and expertise between the institutions.

JAMES ALTUCHER: Interesting. Cornell in Manhattan, again, I'm talking about 25 years ago, used to have a computational finance... Their computational finance department was in Manhattan. I don't think it exists anymore there, but I'm not sure. But anyway, thanks so much for coming on the show. This is my favorite topic. The Age of Prediction, thank you so much. I hope you guys come on again, I really appreciate it.