MILKEN INSTITUTE

CONVERSATIONS WITH MIKE MILKEN

James Golden

CEO, WorldQuant Predictive

May 12, 2020

Mike Milken: Jim, thank you for joining me today.

James Golden: Thank you, Mike. It's a pleasure to be here.

Jim, many people might not be familiar with the company you lead, WorldQuant Predictive. So I thought we might start by discussing why it was created and what kind of work it does, and later we'll discuss its relationship to COVID-19.

That sounds great. At our core, we're a data science company that leverages a combination of AI, machine learning and quantitative methods to help solve really tough challenges. And we believe what's unique about us is that we have developed approaches to making predictions at scale based on established principles that have been successfully used for many years in quantitative finance. As you know, many of these methods were pioneered by Igor Tulchinsky, who's the founder and CEO of WorldQuant Asset Management. What we do, our team leverages these approaches to solve predictive challenges that exist today across a range of industries, including things like COVID, which we'll talk about.

This interview has been lightly edited for clarity and readability.

As you well know, WorldQuant isn't the only firm constructing complex algorithmic approaches to investing. There are many brilliant, successful quant finance companies. But what's really unique about Igor's approach is that he's been scaling up the number of signals that can be found in multiple diverse datasets, and he assembles these into what he calls an "Alpha Factory," first generating dozens, then hundreds, then thousands of algorithms which can be packaged into large portfolios for trading financial instruments.

"When the Black Death swept across Europe and wiped out a third of its population, it dismantled feudalism. … Spanish flu alongside World War I led to mass urbanization of the U.S. … What can we learn from the past that at least gives us hypotheses that we can test in models to see what's going to happen next?"

The real advantage, I think, of having so many models used to make predictions is just the power of competition to produce winners. Igor's Alpha Factory was designed to be flexible and expedient, and it operates by aggregating many signals together to create models with greater predictive power. Now, coming back to our company, Igor soon realized that these ideas had applicability to many other areas outside of finance. So we set up a new company two years ago called WorldQuant Predictive to apply these iterative ideating, testing, competing approaches to problem solving to build what Igor calls a "Prediction Factory."

I think what's most interesting for me is I remember the moment when I fully grasped what WorldQuant's approach to prediction was really about. I was visiting different WorldQuant offices around the world, and the head of one of the offices mentioned to me that one of his researchers had come out on top for several months among which researchers predicted the best models that generated alpha for the hedge fund in the market. I met this young researcher standing at the coffee machine and I flippantly asked him, hey, what's your secret? He looked at me and he very casually answered, "every day I walk a different way to work." I was completely blown away. That was my Eureka moment.

If you're going to accurately predict, you require algorithms that approach the problem from many, many different directions – the more the better. To do that, you need model builders who think creatively, with diverse datasets that offer as many different perspectives as possible on whatever problem you're trying to tackle. That could be from trying to predict the path of a stock, to attempting to discover why people are doing things like hoarding toilet paper during a pandemic. Scale is essential. Diversity is very essential. The other really nice thing about this approach is this helps remove bias from AI systems, one of the biggest challenges. The greater adoption of AI, diversity of people, data, ideas – it reduces bias in hypothesis generation and prediction.

So Jim, as you know, I've known Igor and we've been friends for a long period of time. We both have a love of mathematics. One of the things that really was eye-opening for me as I looked at this was the idea that Igor had put forth that many others have discussed since, is that intellect is evenly distributed throughout the world, and by getting people throughout the world to work from their perspectives, you actually get better results.

What is your background, and how does that fit with WorldQuant's mission?

My career has mostly existed at the intersection of data science and in healthcare. I started out studying mathematics and computer science in college, and then I spent a number of years actually in the U.S. Air Force, and I was very fortunate to find myself at the Air Force Test Pilot School in California, up at Edwards Air Force Base, teaching classes in mathematics and flight dynamics to an amazing group of individuals, men and women, many of whom later went on to become NASA astronauts. That got me very interested in engineering. So I went back to graduate school. I went to Vanderbilt to study mechanical engineering.

At that time I was dating a young lady in the med school and she had just started working on DNA sequencing, which was very nascent at the time. And so the Human Genome Project had just gotten started and she had one of the very first DNA sequencing devices. It was very primitive. It didn't work all that well. Four-colored DNA sequencing had just been invented. So I took a look at the device, re-engineered a few parts, and started writing some software using early neural networks to better read and interpret the signals from the device. That really led me to my passion for machine learning and artificial intelligence. At that time, the early 1990s, biomedical engineering didn't really yet exist as a field and bioinformatics as a discipline was just getting established.

"When we think about data, we have … EMR data, genomic data, sequencing data – all of those things are extremely valuable and tell us a lot about viral mutation, virulence, what happens with comorbidities. But there are other kinds of data: mobile phone data, transportation data, consumer signals, buying and demand curves."

But the really pivotal moment for me was right after I finished my graduate work. My mother was diagnosed with a glioblastoma – a disease I'd never even heard of. When I asked my family physician what we needed to do, his response was pray that she goes quickly. I was absolutely outraged by that, and it completely changed the course of my career. So I went into biotech. I was trying to figure out how AI could be used to interpret all the data that was coming out of the Human Genome Project, especially in the areas of oncology. Later I started thinking about how all of that genomic data tied back to real-world clinical data and what that could mean for patients. I did a lot of work on IBM Watson at places like MD Anderson and the Mayo Clinic. And in 2018 I actually met Igor Tulchinsky at one of your global meetings in LA.

Jim, you've touched on a number of things. Like you, I've been thrust into focusing on glioblastoma – in the 1970s with one of my only two first-cousins, to my stepfather, to my close friend Reginald Lewis (whom I had a chance to finance and to me is the Jackie Robinson of business for African Americans), and lastly my sister-in-law. In fact, the March we put on in '98, which culminated the work of thousands – we had a half a million people in Washington and around the country and spent three years planning it. A month or so later president Clinton signed into law the doubling of the National Institutes of Health budget, the tripling of the National Cancer Institute and many others. But I had met a mother that day whose second daughter had died of a glioblastoma, and her daughter was fighting to stay alive to join us at this March for almost a year. And she died the day before. Yet her mother came. I was just thinking about the family and she told me that her daughter had died more than a decade after her first daughter and they gave her second daughter the same treatment as the first daughter, even though it didn't work 10 years before. I think that's why that doctor told you that. But there are some breakthroughs today and much of our work in immunology is now going into glioblastoma patients. But this brings us to COVID-19.

Could we have predicted COVID-19? Could your models have predicted it? How do you reflect on that today?

I think to a degree that we knew that some kind of pandemic has always been inevitable. We have many examples: Ebola, H1N1, SARS. We've been talking about these things for a long time. What we didn't really predict was the magnitude of the impact, the stress on the healthcare system, shocks to markets and economies, impacts from social distancing, shelter in place rules, consumer hoarding behaviors.

If we go back – and I always go back to that idea about walking a different way to work – this is where we have to start thinking about our world in relationship to the COVID pandemic. Our world has changed, and we are all going to have to start walking a different way to work. We need to ask, what does that mean for us? How do we predict the new normal at speed and at scale? We have some really interesting examples, historic examples. The plague – when the Black Death swept across Europe and wiped out a third of its population, it dismantled feudalism. It completely changed sources of governments and the land that had been the primary source of wealth was now worthless. Spanish flu alongside World War I led to mass urbanization of the U.S. Individuals giving up the farm, going to work in factories. What can we learn from the past that at least gives us hypotheses that we can test in models to see what's going to happen next? For us, the way we think about that – any problem, not just a pandemic, but how do we respond to any natural disaster – is to look at data, find signals in that data, build models and make predictions.

Jim, as I look at data, I step back today and say, South Korea is back to work – most of the people are in their offices. Taiwan is back to work. China is back to work. Thailand really never stopped. Many of those had the experience of SARS, H1N1 and other types of things. And as we step back today and look at Europe and the United States, there was a lot to learn.

Let's talk about these predictive models. Hospital providers have lost billions and billions of dollars by committing themselves and numerous academic centers for a potential surge of patients that never came. So their occupancy is low. If we look at the National Cancer Institute's own hospital – and most of the people that have come there are headed toward end stage or death that were not able to be saved at other locations around the world – it's operating as Steve Rosenberg pointed out in a podcast at maybe 10% of occupancy.

It isn't just the prediction of what might happen to get prepared for, but it's also the prediction of what the needs are and what the demand is so that we're properly prepared. Can we do that and predict the next pandemic? Can we look today and say what's going to happen in the fall? Are we going to have a resurgence?

I think we can. The key is, how do we connect those dots to establish predictive models that really allow us to take action? But also for those models to evolve over time as we get more data? One of the most interesting things when I've talked to my colleagues in medicine is why are there so many differences across the globe? When we look at the cultures of South Korea, Taiwan, China, and others, there's much less handshaking or hugging. There's a lot more bowing. Is that possibly a way of reducing the spread of disease? When we look at countries like Italy, we have multiple generations living in the same household. Is that something that helps spread the disease in that region of Northern Italy where the virus has been so destructive?

When we think about data, we have all kinds of data. EMR data, genomic data, sequencing data – all of those things are extremely valuable and tell

"Our retail clients are asking us questions like, will people stay at home this summer? Even if we lift shelter-in-place, is this a summer of backyard family barbecues, and what does that mean for our inventory? Should we stock more lawn furniture, garden tools and barbecue grills?"

us a lot about viral mutation, virulence, what happens with comorbidities. But there are other kinds of data: mobile phone data, transportation data, consumer signals, buying and demand curves. How are people behaving? We can tell if people are standing close together. Are they buying things or are they waiting to pick things up or are they moving in and out of a store? What does that mean for, say, ventilation in storefronts? How many people should be allowed into a store at any given moment? We know that everything is going to change, especially around things like retail. Are people going to travel? What does this mean for the airline industry? Are middle seats going to be banned, which I'm sure many people would be happy about, but what does that mean for the airline industry itself?

"When we look at the cultures of South Korea, Taiwan, China, and others, there's much less handshaking or hugging. There's a lot more bowing. Is that possibly a way of reducing the spread of disease? When we look at countries like Italy, we have multiple generations living in the same household. Is that something that helps spread the disease in that region of Northern Italy where the virus has been so destructive?"

So as we think about those big connected pictures of how we go forward, can we create better healthcare models? We look at things like influenza models – and we've helped some of our clients understand prediction of infectious disease. But when we start layering on things like signals from transportation data, say airplanes that take off from the city with higher rates of infection flying to a city with lower rates of infection, what does that mean for a vector? Will we see an increase, when we think about weather data, where there's some places are warmer or chillier than others? Does that actually help improve the accuracy and the resiliency of those models?

This can really help us understand things not just about, say, how the disease is evolving, but how far behind in New York City are we? And that can help us think about how we adjust behaviors, resources, and responses accordingly when we create these more rigorous benchmarks.

One of the things I have to assume is you're constantly being asked now by clients. What are retail clients asking? What are real estate clients asking? What are healthcare clients asking?

Sure. And they're all asking a variation on the same question, right? So again, when you and I met for the first time, the question you asked me was, what is the most relevant pressing question that every business needs to know? And we now know that is, what is the new normal going to look like post COVID?

Our retail clients are asking us questions like, will people stay at home this summer? Even if we lift shelter-in-place, is this a summer of backyard family barbecues, and what does that mean for our inventory? Should we stock more lawn furniture, garden tools and barbecue grills? People are also asking us, especially different retail and consumer packaged goods companies, are saying, hey, we know people have been creating hoarding behaviors and they're putting away toilet paper, pasta and other things. When things get better, should we begin manufacturing those materials again? Or is there still enough being hoarded that people are not going to change their buying behaviors for some time going forward? So those are really kind of interesting things about what does that mean for manufacturing? And obviously all these things trickle down to other industries. It's about supply chain, it's about pricing, it's about consumer demand. Can we build models that are active, not brittle, that continually evolve in the face of new data?

We've started doing a really interesting project working with the American Chemical Society. Here's a great example of looking at alternative data to create a different approach to respond to COVID. The American Chemical Society has a group called the Chemical Abstract Society, which is sort of the keeper of all information about all known chemical compounds in the world and they have a great resource around medicinal chemistry.

"My mother was diagnosed with a glioblastoma – a disease I'd never even heard of. When I asked my family physician what we needed to do, his response was pray that she goes quickly. I was absolutely outraged by that, and it completely changed the course of my career. So I went into biotech."

They've asked us to help them provide a curated resource to assist drug discovery and development for researchers in the hunt for new COVID- 19 therapeutics. They've created a library of roughly 50,000 known antiviral compounds, and we're providing AI-based predictive analytics to help them create new connections and new associations around this data to answer several relevant biological questions. Clearly people are looking at things like cytokine storm, which you've talked about before. When you see a cytokine immune response, how does that trigger biology where we can actually design a new compound to match against that?

When we look at lungs that are filling with fluid because of an overabundance of inflammatory response, are there chemical compounds within that database that could be matched against those target pathways? What we're going to do is take that compound database, tie that back to all of the biological information that we're currently gathering from places like our friends at Weill Cornell hospital, and then we're going to release that into the public domain so that pharmaceutical companies can actually prioritize some of their research. Obviously this is very tied back to the wonderful work that you're doing at *FasterCures*. I've talked to Esther about this a little bit. We're really, really excited about taking what we've done and tying that back to your work and tracking different compounds and therapeutics and vaccines.

Predictive models can tell us what's going to happen, what should happen, but it's the innovators who are attempting to change the course of history. As you know, we are now following more than 200 different applications in vaccines, in antibody work, antiviral work, immunology work, and many of these are moving along. The federal government led by BARDA has agreed to finance the production of vaccines before we know whether they work. And if they work, they'll be available. In the case of large pharma, they've made contributions. In the case of biotech, BARDA and others have supported that – Wellcome trust, our own foundations, the Gates foundation.

Part of our effort, Jim, in many ways is to change what the predicted future would be. Can you change the course of history? In a financial crisis, many companies might fail, but can you create this financial safety net that we've been working on with the government, economic and financial, so they don't fail and they can make it over this difficult period.

Let me take you to this next level and next challenge for WorldQuant Predictive. So all this analysis, your brilliant people in more than 20 countries all over the world – as you're servicing clients and telling them based on this data, this is what is going to happen. But what about the data, Jim, that appears to show that if you don't have preconditions and you're not a senior citizen, and we've been analyzing why senior citizens, why age is such a factor that for the most cases the virus does not cause as much damage. You have to be able to analyze this data and ask, should we move to a policy like Sweden to let those that are not at risk return to normal, while those that have preconditions or are older or very young (under 1 year old) remain sequestered for a longer period of time?

What you're really talking about, Mike, is in our culture, in the fabric of our society, how do we think about resiliency and social cohesion in the face of a dramatic crisis like the COVID pandemic? Where does prediction data model building allow us to really get to the point of what do we need to do that doesn't just mitigate spread of disease and eventual mortality, but economic survival, survival of social institutions. That really gets back to the heart of what we do, which is taking a multitude of different ideas where we ideate, arbitrate and then make predictions on that.

We have to look at places like Sweden where they're trying a different experiments. But we're seeing the rise of deaths in Sweden to be much higher than that of surrounding countries. But what does that mean? Does that mean we try this as a great experiment and accept the risks? Not everybody will do that. Really when we think about resiliency and cohesion, how do we find data sources so that when we put a policy in place we're getting immediate feedback on not only how that policy is working for, say, job disparity, income, payroll, but where we're seeing changes in behaviors like social distancing and how those are influencing things like the spread of the actual disease. It's a really difficult challenge, but I think it's one where we actually have the ability to create new kinds of methods to really predict that new normal in real-time and respond accordingly in a way that limits the loss of life, that increases our ability to give the best kind of healthcare to everyone in the United States, but to also be able to say how this affects economic security and mitigate those underlying effects to society as a whole.

In 1974 during an epidemic like this in the financial markets, I was on a panel where one of the leading academics had predicted that 700 of the 2000 largest companies in America would go bankrupt, and I suggested that people will not stand by for Lockheed to go bankrupt. People will step in and it won't happen. I was on the same institutional panel in ’77 when this individual stated that he had predicted four of the major bankruptcies that occurred during this three years. What he didn't tell people is he had predicted 700 to go bankrupt. My point was ConEd didn't go bankrupt in New York state, and Lockheed and all these other companies did not go bankrupt because people took actions, changed their financial structure, and other activities that occurred. So it didn't happen.

I think this issue is the challenge I'd like to give you when we do our next podcast. You might be better able to predict what's going to happen than anyone in the world today. Let's see if we can take the next step with the people you interact with here. With that team, let's change the number of people who are going to die, the number of businesses that are going to be disintermediated or go out of business, the number of lives that have been devastated, et cetera by using your predictions. Let’s start with what you're doing with all the chemicals that might address and attack the virus and kill it or prevent the cytokine storm. Sounds like a good way to go.

Good health to you and your family and your team, and we look forward to you changing the course of history.

Thank you, Mike. Real pleasure talking to you today. Thank you for all you do.