NIIT Technologies' Sudhir Chaturvedi: Companies Have to Decide What They Need from Big Data

Big data and predictive analytics are creating a big buzz across industries. They are seen as critical tools for making fast and accurate business decisions and a huge source of competitive advantage. In this interview with Knowledge@Wharton*, Sudhir Chaturvedi, chief operating officer of NIIT Technologies, suggests that to derive maximum benefit from these new technologies companies need to first define the business outcome that they want. "Define that in a granular manner and then go after it and use data to see how you can make better decisions in the market," says Chaturvedi.*

An edited transcript of the interview follows.

Knowledge@Wharton: There's a lot of buzz today about big data and predictive analytics. Could we begin by speaking about what these terms mean and some of the reasons why they have become so popular?

Sudhir Chaturvedi: You have probably read that 2.5 quintillion bytes of data are created every day. That's 2.5 exabytes of data — a lot of information. This information is in very different formats. Corporates refer to it as internal and external data or structured and unstructured data. Today, the sources of data are several – sensors, GPS signals... The sources of data have expanded exponentially. When you put this data together, it leads to data which we call big data, mostly because of the quantum of the data, the velocity at which this data is created and the variety in which this data exists.

When it comes to predictive analytics, that is a bit simpler to explain. I'll use a Wayne Gretkzy [former Canadian ice hockey player] quote to explain this. Gretkzy said: "I skate to where the puck is going to be rather than where it has been." And that's really predictive analytics. Don't tell me what has already happened and why it happened; tell me what is going to happen and what should I do about it. These two topics sit very well together because predictive analytics relies on big data to be able to make better and more accurate predictions about where businesses should make certain decisions.

Knowledge@Wharton: So, building on that, if you were to look at how companies in different industries are using predictive analytics, could you explain how they are integrating data that is generated both inside and outside the enterprise to do things that you couldn't in the past?

Chaturvedi: Almost every industry now understands the importance of big data and predictive analytics and they are spending a lot of time and effort on this. I think that some of them are doing a great job, but the majority has some way to go. I hear my clients telling me things like: "The quality of my data is not good. It's in too many places. We have too much duplication. We have too many owners of data. There are too many silos." What this engenders is a bit of mistrust of data. This is interesting because we talk increasingly about decision-driven organizations in the world of business and decision-driven organizations led by data. But organizations seem to have an issue with the data that already exists within the enterprise.

There are good reasons for these misgivings. They are not wrong. But the point is: what do you do about it, especially when you add a whole load of external data from social media? I don't know if people realize it, but every time somebody tweets something, they are actually sending 21 different pieces of information in one tweet. How do you integrate that information with what you already have?

When you overload with sensors and data you are getting from other sources, it becomes quite a complex thing. That's to sort of paint a picture of the complexity and why people have found it difficult to really make big strides.

Technology today is much better than it used to be a few years ago. Actually, technology's ability to deal with the complexity is great. We can create data structures that are easy to use. We can run algorithms on it which are fairly simple to use. But the issue that remains is, business trying to figure out: "What can this do for me?"

I recently met a client and he told me: "I have all this data, why doesn't it tell me something?" That is a red flag for me because it's like a TV crime drama quote: "You torture anything enough, it'll confess." But will it really confess what you want it to confess? That's a question that can equally be asked of data. So, what's most important for corporations is to actually try and figure out what is it that they want from that data. First, define the actionable business outcome that they want. That needs to be defined at the level of granularity so that a person who is in the field and needs to make a decision based on that predictive insight is able to so. That's what I call an actionable business outcome.

Corporations, in order to use big data better, need to think small instead of thinking big. They need to think granular in terms of what they are really looking to do. Even if they do it in four or five instances and they do it really well in terms of actionable business outcomes, they're going to get much more value out of it than they would if they just looked at the entire quantum of data and wondered, "What can I do from this?"

Knowledge@Wharton: Could we take a couple of examples from different industries and see how this plays out with them?

Chaturvedi: One of the best examples I have seen is in the insurance business. The insurance industry has more data than you can shake a stick at. But it has traditionally been a little circumspect in terms of how they have used the data. But, today, we're working with clients in, say, home insurance. All you need to do is enter your post code and your home number; it gives you an instant quote. You don't need to fill in any data there.

Is the insurance company taking a risk? They're not because the amount of data that is available from a post code perspective — which gives you a very good idea about the risks you carry — is tremendous. It's about whether you have the ability to put that together and present it in such a powerful way that someone enters a post code and within nanoseconds they're getting back a quote. Later, of course, there is a process to go through, especially if there are claims outstanding or things like that where you would refine the quote. But, in general, these clients are also offering a guarantee that their final quote will be within 10% of the quick quote that they're offering.

This is in the consumer experience. But I'm also saying this in the commercial experience where the stakes are much higher. In the commercial space, the level of information is much more. If you take all the data that's already available around insurance in different markets — which building has been insured at what price at what point in time — and overlay that with ready data that's available or data that's available from Google which is integrated with some of the things that we do around geo-coding, around street view etc., suddenly you find that there is a lot more information available where you can

marry structured and unstructured data together and provide a much better service to a client because a client can then make quicker decisions.

This is in the insurance industry. And I can sort of give you several examples of where great things are happening. But one of the big things that I see is: "How do you account for customer churn? How do you predict customer churn?" Because a lot of churn happens without a lot of people even realizing that it's happened. And it's one of those things that every one of us knows. It's sort of seven times more expensive to get a new client than to keep an existing one. Yet you see almost every industry spends all its efforts getting new clients rather than keeping existing ones and is not very good at predicting which of the existing ones is going to leave.

This whole thing about churn prediction — today, again, once you marry the structured data that is already within the enterprise with data from outside and you put some algorithms around human behavior which are well proven now, you can actually start to predict churn in a very good way. You can look at it from two perspectives. You can either look at it from a competitive perspective where you can predict somebody else's churn and use that as a lead generation engine for yourself to see how you can get more customers. Or you can use it within the enterprise to say: "How do I reduce my churn?"

For example, we had a wealth manager who was looking at college fund investments. There is a very finite time in which this investment is going to be used and you're worried that the churn might happen in the middle because of performance. Now, that's a very specific case. Now you can predict churn more accurately in those kinds of cases rather than predicting across a whole universal set of all customers that a bank or an insurance company might have.

So, that's what I meant, going back to the previous point I made. When you think of big data, think small, think granular, find out these kinds of things — business outcomes that you want to influence and then work around them. There are solutions that are available today to help you do the right thing, mostly for your customers. You do the right things by them and eventually you will benefit too.

Knowledge@Wharton: In your experience, when companies try to implement some of these projects, what are some of the common pitfalls?

Chaturvedi: This is an area where companies are traditionally structured around people who do functions — line functions or staff functions. There are people who do the measurements etc. Companies aren't really structured to be data driven. So, companies need to look at how they can work on this in every part of the enterprise. One of the things I would recommend is creating functional teams of people where somebody brings in the deep business expertise that is needed to find out that granular level of business outcome that you want to influence. You need to bring in data scientists — statisticians and mathematicians — to help you with this. You need a technology team to also make sure that this data is available for the mathematicians, statisticians and the business experts to work on it.

I would say that every organization needs to look at creating these cross-functional teams in areas where it wants to influence those kinds of business outcomes. That is an investment that will need to be made. I would not advise having this as a separate function for the entire organization because then you are more likely to get into that thing that I mentioned earlier — "I've got all this data, why isn't it telling me something?"

Knowledge@Wharton: In your own personal experience, what instances have you seen of value generated by judicious use of big data and predictive analytics?

Chaturvedi: Going back to the example that I talked about — quick quotes — this is something that we are doing for one of our clients. It's been up for about four months and that's already resulted in 10% more bookings than before.

Another example is our insurance products business. One of the things we have seen in insurance is that as investment incomes have come down because of interest rates being fairly low, the need to make an underwriting profit has become more paramount. Now, you would imagine that most insurance companies exist to make an underwriting profit. You'd be surprised how many don't make a significant underwriting profit. But the thing in underwriting is that previously underwriters looked at maybe 25, 30 variables to come up with underwriting decisions on insurance. Now, you can in theory look at an infinite number of variables to come up with the same. What we are saying is that judicious use can lead to a 10% reduction in the loss ratio.

That is very significant. It could mean the difference between making a significant underwriting profit versus a significant loss. But, again, I keep going back to the point that people are choosing very specific areas within their business where they want to influence the outcome and then using the data to arrive at that outcome.

Knowledge@Wharton: If we were to look at a crystal ball and look at the future, where do you think this whole field of predictive analytics and big data is going to go?

Chaturvedi: Well, I think it's only going to get bigger, especially now that we have 2.5 quintillion bytes of data already available. As the Internet of things become a reality and devices start to talk to each other, you will see an explosion of multiple forms of data coming into the enterprise. It can be a huge source of competitive advantage. You don't need to worry about the technology complexity; that is actually the easiest part. What you really need to focus on as an organization is: "What is the outcome that I want? What is my business strategy predicated on? What are the successful business outcomes I want?" Define that in a granular manner and then go after it and use data to see how you can make better business decisions.