How To Use Data In A Crisis

Business intelligence expert John Parkinson on how leaders today need to prepare for the new environment

Stanton Chase recently convened a group of thought leaders to discuss how new data practices are reshaping the corporate landscape and lighting the way forward for those who innovate wisely. This month, we followed up with business intelligence expert John Parkinson, the former CTO of TransUnion and CapGemini and a leading figure in the field of data innovation, to delve deeper into the question of how companies and business leaders can better use technology to survive and thrive in today's environment.

As a partner and founder of Parkwood Advisors, Parkinson provides intelligent data solutions for clients like Amazon and the World Economic Forum. Throughout our interview, led by Greg Selker of Stanton Chase, Parkinson shared his insights about the novel ways that data is being used to respond to crises — and the common mistakes and false assumptions he has seen companies make as they strive to adapt. We are delighted to share the full version of this timely conversation.

Greg Selker: When you look at the state of the world today, what are the data-oriented questions that you believe businesses aren't thinking about but should? How should companies seeking to adapt to this new environment think about data?

John Parkinson: When you're in a situation you have never encountered, you don't know what's important so there is a tendency to chase things that look interesting, whether or not they turn out to be important. Today, that tendency is exacerbated because we have lots of data. Our customers expect definitive answers that we can't give them. We can't give them the "right" answer because there isn't one. Trend analyses and predictive analytics really require some sense of the structure of causality that you're looking at, "if this, then that." We've stopped trying to offer this type of causal prediction. Instead, what we're saying to clients is, "tell me what you want the outcome to be, and we'll tell you what variables to watch to see if your desired outcome is going to happen."

JP: We try and run as many scenarios as we can. If I have a model that I think is predictive of the outcome that a customer wants, I'll run a sensitivity analysis. I'll run 100 million scenarios until I figure out what factors change the answer because those tend to be the things that you should be paying attention to.

GS: Can you give an example?

JP: We recently had a client ask us, "As we look at coming out of our current crisis, should we keep our current office footprint or continue to have the majority of our employees work virtually?" They have 40,000 employees today and, currently, about 1,000 of them are working from company-owned real estate. Right now, most of their office space is empty. That's

expensive overhead. So, the client wanted to know if they should downsize or relocate, and they also needed to consider how those types of changes to the office footprint would impact the number of employees likely to return.

To explore a set of questions like this, we build models. We mine all the proprietary and public data sources we can and take hundreds of factors into consideration. It gets complicated pretty quickly. For instance, if we know that most of their employees drove to work prior to the pandemic, that means that most employees will not be coming into contact with anyone during their commute, so we can make that a low-weighted factor. However, when we look at the office itself, many of the cubicles are only three or four feet apart, which makes social distancing difficult, so that factor may be more significant to the outcome. The model also considers factors like the age demographic of the workforce. We can make assumptions about the propensity of people that want to work close to other people if they're 25 versus if they're 50. We also look at the emotional response of their workforce in terms of COVID-19. How safe will people feel? We look at and weigh hundreds of factors.

Once the model is built, we run lots of iterations to see what changes. It's cheap to run; we did about 1 million runs in this particular example. I think it cost $25.

GS: I'm struck by that. Ten years ago, instead of a day to run that kind of model, I think it would have been days or weeks — and at what expense?

JP: Right. Cheap computing has made a much bigger difference than most people realize. Ten years ago, when I was at TransUnion, running models like this could take weeks. I think our record was a model that took six weeks to run; it ran on 8,000 Intel cores in a data center that we owned. That's about $2.5 million worth of hardware, plus a whole bunch of people to look after it. Today, all I do is load it up in a bunch of containers on AWS and just press the go button.

GS: It seems like, post-Covid-19, emotions like fear are having a greater impact on how people move in the world. Is emotion somehow captured in your models?

JP: We build in some variables that have to do with people's behaviors. Our baseline variable is that people will feel just the way they did prior to Covid. Then for each demographic or each cohort that we've modeled, we add in some variables that say, "That will change." Perhaps they will be less likely to want to get together because they're concerned about community transmission and getting sick; or perhaps they will want to get together more than before because they're depressed and haven't seen anybody in months.

One of the hard things about giving people definitive outcomes from this is we're saying to them, "Look, we can have people who are scared but decide they need to do this. And we have people who are scared and won't do it, and we have people who are not scared and won't do it." So, remember: We're not trying to give you the answer; we're trying to tell you what to look for to figure out whether the outcome you want is going to happen.

GS: Are businesses today taking advantage of computing power and strategic computational models to guide their response to COVID-19? If not, what do they need to do to take better advantage of the technology that we have accessible to us today?

JP: Every model you build is a simplification of reality, and your model's usefulness is as much defined by what you throw away as what you keep. So although the cost of computation is very low, there is a significant cost to collecting, curating, and using the data.

Historically, businesses tended to collect and use only the data that they thought was relevant to their operations. They did not generally build these kinds of strategic, analytic models. There are a couple of reasons for this: one is that the talent component of this does not come for nothing — the talent to do this type of data modeling is relatively rare. The other reason is, over the years, there has been a tension between investing in a small group of talented data workers and the democratization of data and analytics, which would equip every employee with the tools to ask more intelligent questions and get better answers.

Unfortunately, we have not been able to give the average knowledge-worker sufficiently intuitive and easy-to-use tools that make asking good questions easier. It turns out that most people don't want to ask better questions. They want to be told better answers.

GS: And why is that?

JP: Personally, I think it’s human nature. Asking good questions is actually hard work. So, we have historically not been able to give the average knowledge worker sufficiently intuitive and easy-to-use tools to allow asking good questions to become easier.

GS: So regardless of the fact that we've got more data at our disposal and greater capabilities for quickly analyzing that data, it still comes down to talent. There is still only a small group of people who truly understand how to pose the right questions, curate the data, and construct useful models. This is a highly valuable and relatively rare talent among all knowledge and data workers.

JP: Yes. I think that's one factor. The second issue is the investment return issue. I'll give you an example. We had a customer, a big retailer, who told us that they had so much data that if they knew who you were when you walked in the door of one of our stores, they could predict with more than 90% accuracy what you were going to buy. However, their physical supply chain operating capability could not respond fast enough to the fact that they knew what a customer was going to buy. So, they could improve their predictions of whether a person would show up in one of their stores, and once there, what that person would want to buy, but even with that capability, they were still going to be wrong often enough that it sort of wasn't worth the investment to try to improve their supply chain to be more responsive to these capabilities.

Their CMO said to me at the time, he said, "We can change our inventory in about four months. You can change your mind in four minutes."

So that's a factor. It's a question that we always ask customers who approach us with that kind of problem: "Are you willing to make significant investments to your overall operational processes as a result of what you learn from this exercise?"

GS: So in this example, you're asking, is this company really willing to invest in changes, not only to supply chain management, but also to overall IT infrastructure, logistics, transportation, and vendors, etc, in order to drive maximum efficiencies into their business?

JP: Yes, exactly.

GS: Even if a business today acquires the right talent to take advantage of the cheap and speedy computational processes that are available, there is an additional barrier to success. To really drive the conclusions, leaders need to be ready and willing to make significant investments— not just singular investments but expansive investments that cut across multiple areas of a company's operations and infrastructure.

JP: Yes. So you have just invented robotic process automation, or RPA.

GS: Well, I don't think I invented it, but I might have articulated it.

JP: What do we do in first-world economies when we don't have enough talent to solve a problem? We turn it into software. So now, though RPA doesn't necessarily ask the best questions, it asks better questions most of the time.

GS: Can you give an example of that?

JP: Think of customer service. I need to make a decision about whether a given RMA request should be accepted or not. With RPA, I look at who the customer is and how much business I've done with them. I look at the typical return experience for the item in question; if this return is common, I should be more willing; if it's a rare thing, I should be more suspicious. I also look at what the current capacity for my reverse logistics looks like … and I do all that in 100 milliseconds.

No human can do that even if they know the rules. Even if they have a checklist to go through, the human is going to take five minutes. The RPA is going to take 100 milliseconds. The customer gets an instant answer and a rationale, and, in general, their satisfaction scores will

be higher. Plus, I didn't have to have a person involved who might be having a bad day, or tired, or not feeling very well, or doing this for the third time in the last 20 minutes.

So in order to solve the, "I don't have enough good questions or talent" problem, we create an automated process that means I don't have to have as much talent. Now if I'm smart with how I design the data-driven RPA, I funnel the ones that the automation can't handle to a smaller group of more talented people to deal with the edge cases.

GS: So ultimately, even if you have successfully implemented an RPA, the question remains: Are companies willing to take the necessary steps and drive the necessary actions across multiple operational areas to implement the suggested improvements of an RPA?

JP: Yes. But now we should circle back to the question that was originally asked by the model, because the question in our example was, "Are these employees going to come back?" So, in this case, part of the model output is, "If you no longer need them, why would they come back?"

If you're going to make the investments in having a higher level of automation in your business processes, one of the outcomes is that you will likely need less people to run them, but the people you do need will have to be more talented; therefore, you probably need to be thinking about how to encourage that small group of talented people to return. As I said, we're not giving people answers, we're telling them what to pay attention to.

GS: In your model, does the length of time people are hunkering down due to Covid drive whether it's a more permanent change in behavior?

JP: Oh, that's a really, really great question. The answer is we don't know. That's one of the things that we build watch points for. I have a minor in behavioral psychology and, in that field, there's intense debate over whether prolonged exposure to a new circumstance makes you more or less likely to want to persist in behaviors you've adopted in response to these circumstances. As usual in behavioral psychology, the data says, "Yes, that is correct. Some people will persist in these new behaviors and some will revert to old behaviors." But it won't tell you for any given individual which way they will go. However, you can build statistical models to say that some people will get to like it and want to stay and some won't. Some people will say, "I hate this stuff. I haven't seen a human in weeks. I'm tired of wearing a mask all the time. So, let me go back, even if back is different. I don't want to stay where I am." Meanwhile, some people will enjoy working from home, never interacting with coworkers, and avoiding their hours' long commute. You're almost certainly going to see both outcomes.

GS: Well, isn't getting that question answered really impactful in answering the question of what the future of retail is?

JP: Yes. Absolutely it is. And what do you do? You build scenarios and you ask, “How do we design a data environment in which we can see which outcome is happening?” Supposing the answer is “both of the above,” then how should we plan to position ourselves so that we can serve both the people who come back and the people who don't?

GS: Right. What do we need to do so that we can best survive and actually thrive with this reality, this duality, in which some people — both customers and employees — are more comfortable interacting with us virtually and others want to interact with us physically and not just in a virtual fashion.

JP: Yeah. Omni-channel was a word that most people couldn't spell, let alone understand, and now it has become the probable reality that everybody has to deal with.

GS: It’s interesting the way in which the definition of “omni-channel” has evolved. Originally it referred to the various channels of a company were deploying to get their message out to a target audience. Now, the word is used far beyond an outbound marketing context. It is really has become an all-pervasive way of thinking about operations. Every interaction, now is omnichannel, as opposed to just singular channel.

JP: Yes, and in both directions and both digital and physical. So that's much more complicated. Thinking about your company in this way logically leads you to say, "I need to know who my customers are." It doesn't matter that I'm a little one-stop retailer. Now I have to be as good at the CRM side of the world as Amazon.

GS: Yeah. And it's not just retail by the way.

JP: Oh, it's everything.

GS: So basically the bottom line is those who guess best win.

JP: That has always been the case.

GS: Yeah. I would agree with that. But if you're willing to take action on what the analysis of the data is delivering to you, then regardless of what the new reality looks like, you'll be in a better position to thrive in that new reality.

JP: That's right. So, you focus on trying to figure out what you have to be able to do no matter what the outcome is. Then on top of that, you put what we call “watch points” in that try and make you a better guesser of what is actually going to happen. Then you need to invest in the ability to be agile in response.

GS: If we look very specifically at the COVID-19 pandemic, what are your observations about vaccine development and the overall virus response? What mistakes or successes are taking place today from a scientific or a political perspective? I know you are not a virologist, but I’m curious to hear your perspective on this as an expert in data modeling.

JP: So, it is amazing that we had a completely tight genome for a novel virus within almost two weeks of getting samples. We've never been able to do that before. That gave the pharmacy industry data to work with to figure out what kind of response the industry could hope to produce. Had we not had that, it would have been trial and error. Computational chemistry allows us to look at the structure of the surface of the virus genome and figure out where there might be attachment points that you could block the virus's action.

However, all this tells us nothing about the important aspects of the impact of the virus on humanity as a whole. We do not have a good model for community spread, we do not have a good model for how the infection actually works, we do not have a good model for the range of impacts on humans. We still don't understand exactly why some people get sick, and some people don't, and it will be likely years before system biology catches up with the reality of what we are experiencing.

GS: Why is that? Why is it going to take years?

JP: Because we don't understand how system biology works. I mean we can statistically predict that some percentage of people with certain genetic traits, blood type, ethnicity, and age will fare worse statistically than another cohort will. We're pretty good at that, but will Joe get it and Mary won't? No one can tell you. We just don't know, and there's a kind of belief in the population unfortunately that science is better than it actually is. Which is not to say that it's not wonderful, it is. It's amazing what we can do, but it's not everything. So, because we have a declining percentage of people with real scientific understanding in the general population, that causes a problem. I have no answer to that by the way. You can't turn everybody into a scientist, but you could at least manage the data in a better way.

One of the things we have been asked a few times is to model worse case scenarios for the virus. The worst-case scenario is we don't get a vaccine and the virus actually that kills most humans. But we aren't looking at that one anymore. Now it seems far more probable that we get a vaccine that's about 60% effective. That would worldwide still probably kill about 1 billion people over two and a half years.

When you do this modeling by the way, which is not a fun game to play, you really do have to look at a macroeconomic impact because it doesn't matter that the virus doesn't kill anybody under 5 because if it kills everybody over 5, all the children die anyway. So, you sort of have to look at population dynamics, and epidemiology and a whole bunch of other factors. The data is all there. The models are not hard to build, they're just gruesomely unpleasant to run. So, you can say, "I have 335 million people in the U.S." You look at the age cohorts that are most at risk, you look at the sociodemographic, and you say, "Well worst case, we could lose 70 million people over the next three years in the U.S."

So that's with a vaccine that's 60% effective. But a vaccine isn't the only answer. What we need is a vaccine and a therapeutic. Now, if we get the therapeutic, actually we do much better. In the U.S., then we probably lose only about 1 million people over the next three years. If you add up all the deaths from the top 10 killers in the population, it's about 1 million people.

GS: And that's just three years.

JP: This is where I go back to my statement, "We don't really understand system biology," for after three years, you've basically eliminated all the people who are seriously at risk. The rest of the population isn't seriously at risk anymore. You have some degree of herd immunity, people get sick, but the healthcare system has adapted to deal with the people who get sick but don't die. Things reset.

GS: And I'm sure included in your model is that fact that we have healthcare workers that are just exhausted.

JP: Yes. Actually, part of the model that reduces the number of projected deaths in the U.S. from 70 million to 1 million is that you have to increase the size of the healthcare worker population by a factor of about five. You end up dumbing down the basic healthcare service provision. You don't need to be registered nurse to look after somebody sick and in the emergency room.

GS: Covid-19, or something similar to Covid-19, was predicted. As we continue to encroach upon the animal kingdom and as viruses become smarter, there will be Covid-25s, or whatever name you want to put on it. There will be more, not less of these issues, coming down the pike in the future. Are we taking data-related lessons from Covid-19 that will better equip us to deal with the next virus? Or do we need to begin more aggressively planning and thinking for that future?

JP: I would say that we're not thinking about it enough. About four or five years ago, we did some work for the World Economic Forum on sustainability, and the question they asked us was an interesting one. It was, "Could you model the natural carrying capacity of the planet?" In other words, without a lot of crowding, huge cities, industrial agriculture, heavy dependence on carbon-based energy, how many people could live here? Like all questions of this nature, there really isn't a definitive answer, but you can model it. The answer we found was in the order of 2 billion people. So, we're at three and a half times that now, going to four times. That doesn't mean that we have to figure out how to get back down to 2 billion. But it does mean that we should understand the consequences of being at 8 billion. There are a lot of things we should be paying more attention to: access to raw materials, waste processing, fresh water, atmospheric pollution, thermal pollution. If you have four times the natural carrying capacity of the planet, you should be engineering things much more than we do currently. We have 265 countries that can't agree on anything. So, on my gold days, I assume that humanity is doomed. But since I probably only have 10 years left anyway, I'm thinking, "Hey, I can probably survive it." But I worry about the world that my daughter will inherit.

GS: Given all of these things, what are the best practices that you would say that businesses should engage in today? Is it to move very quickly into putting RPA systems and infrastructure in place? To move quickly to make certain that you have the right talent? What would be your recommendations?

JP: First, you can't automate everything, so companies should pay attention to where automation is relatively straightforward and reliable. The challenge with automated systems is that when they break, the problem is much worse than it would have been without them because all the routine tasks have been handled. So, by definition, any problem is a bad one. This means that as companies increase automation, they also need to expand their incident response capability to be significantly more sophisticated. Companies should be figuring out how to curate talent, not only acquire it, but nurture it, and encourage it, and invest in it, because you're going to need it. It doesn't matter that you've shuffled off some percentage of your current workforce. The ones you need to keep, you really need to invest in.

Second, companies should focus on understanding their customers. What we say is, "You should not be trying to influence customers. You should be trying to understand them so that the interaction between you is natural, not forced." We have enough data to begin to do that now, although almost nobody does it yet very well.

Finally, if data matters to you, you better make sure that it stays safe, and correct, and therefore you should invest in some capability to curate the data that matters to you. When we talk to business executives, even IT executives, they just don't understand what virtually free compute cycles let you do.

GS: If leaders understood the power of virtually free compute cycles, what changes could that have on what companies are doing?

JP: Well, it means that you can try out a lot of things in the abstract to support your decision-making process. You don't have to be driven by a single decision set anymore. You can play what if and sensitivity gains to a much greater extent than we are used to. A problem is that you give people too many choices, they can't make any—there is a risk of analysis paralysis. So, before you construct a model you also have to ask yourself: is this model actually going to influence me to take different actions?

GS: John, I think that it gets back to talent, but also, the need for an increased investment in RPA systems because the RPA infrastructure will help alleviate some of the potential paralysis by analysis.

JP: It can do. Certainly, the more routine intelligent operations that you can construct in a business, the less things you need to analyze, but we encourage people to model things they are sure about with some randomness injected in because one of the things you can guarantee is you're going to be surprised.

About the Author:

Greg Selker is the North American Sector Leader Software at Stanton Chase International, and a Director in the firm's Baltimore, MD Office.