



WHAT WE MEAN WHEN WE TALK ABOUT BIG DATA

JOSH SULLIVAN works to enable the discovery of knowledge. The Senior Vice President and Chief Data Scientist for management technology and consulting firm Booz Allen Hamilton tells Brunswick about analytics at the speed of business

BIG DATA IS THE BUZZWORD of our age, but if you asked 10 people to define it, you'd probably get as many different answers. The meaning of the term is elastic, partly because the field of data science is so new that even those of us who specialize in it lack clear language to describe its power and potential.

On its face, the term Big Data merely means masses of data, quantities so large that they require new tools for handling and interpretation. It is important to note, however, that the value of all data increases from such analysis – the better the tools, the more useful the data becomes. The real gift of Big Data is in fostering the development of analytics, lifting it to a new degree of sophistication with great promise for the future.

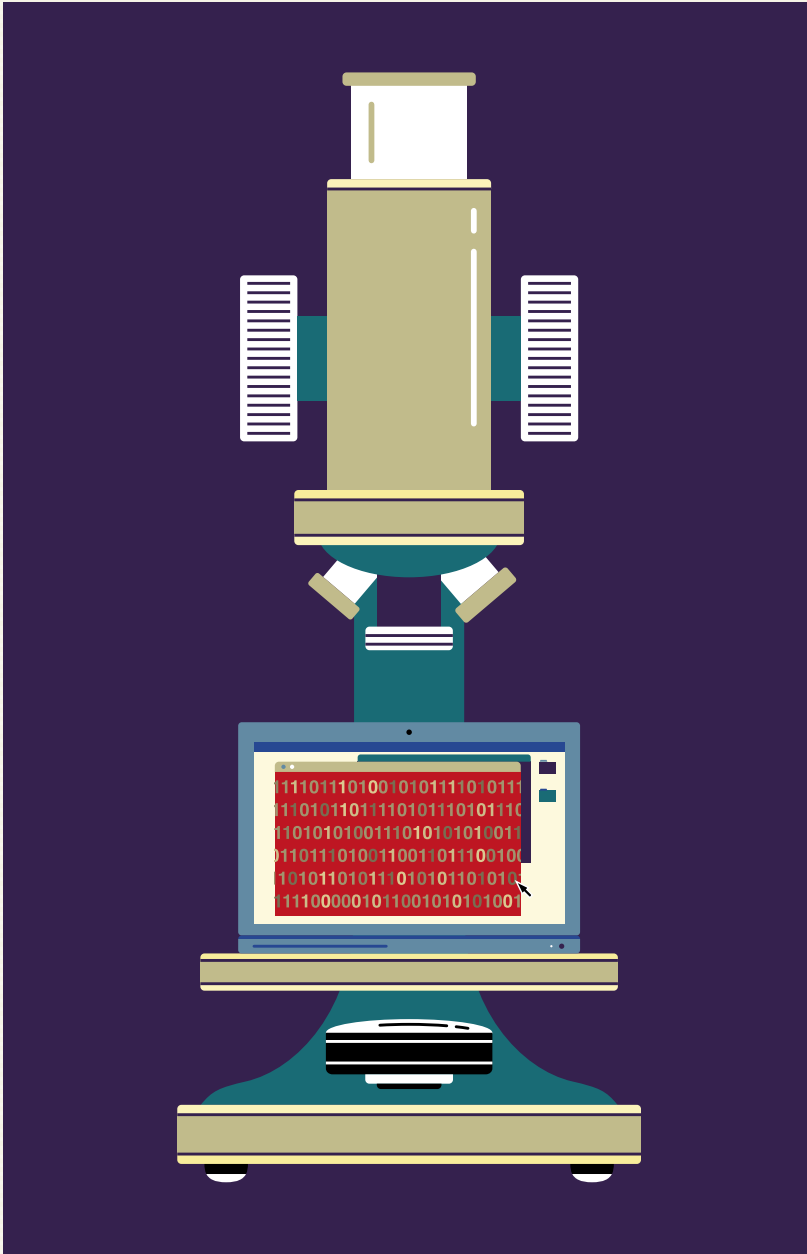
The central premise of data science is that the value of data is unknown until you ask a question of it, test that question, and then deduce an even better, more relevant question. The same pool of data can yield a variety of insights. The goal then, the Holy Grail, isn't just an answer to a specific question, but the process itself, a cycle of experimentation that constantly reexamines data in search of new ways to extract its intrinsic value.

I like to think about this process of data analytics as occupying three areas: *descriptive*, *predictive* and *prescriptive*.

DESCRIPTIVE analytics has been around for a long time. It refers to historical, backward-looking information that companies mine in order to identify past patterns. Historically these have been used to infer insight for the future, such as how to reduce inventory costs or increase asset reliability. People try to guess what's coming by making gut decisions based on what *has already happened* – how much fuel did we use in the fourth quarter last year, and the year before and the year before that? What does that suggest about how much we'll consume this year?

That's a fairly crude approach. The promising area that's developing now is *predictive*, which enables us to do much more complex and nuanced forecasting. This is where, in my mind, the "big" in Big Data starts to come into play.

PREDICTIVE analytics takes the respected, decades-old practice of operations research, a type of forecasting with roots in World War II military planning, and vastly improves on it. Today's scientists are interacting constantly with data, experimenting with different questions and building models to predict what *could happen*. These days, there are ways to run 100 different forecasts every hour, every day.



“The central premise of data science is that the value of data is unknown until you ask a question of it, test that question, and then deduce an even better, more relevant question”

You can ask a new question tomorrow and build a model around it, whether you’re trying to predict cyberintrusions, anticipate failures or forecast the demand for facility space, to give just a few examples. This is analytics at the speed of business.

PRESCRIPTIVE analytics, the third area, is where the value of Big Data gets even, well, bigger.

This entails bringing analytical models and predictions into the real world to help us figure out what *should happen* – the best course of action. While we’re only in the very early stages of prescriptive analytics, it’s where the science of Big Data is headed, and it’s really exciting. I firmly believe that prescriptive analytics will help move us away from gut instincts, enabling companies to make decisions, optimize resources and examine trade-off scenarios in ways that weren’t possible before. In the future, we’ll see data sets on company balance sheets as an asset, as valuable as capital and labor. Predictive and prescriptive analytics are as close to a crystal ball as we’re going to get.

That isn’t meant to imply that the future will be about the rise of the machines. Humans have always had a role and they always will. Machines do analytics, but humans do the analysis and judge what questions should be asked of the crystal ball. Algorithms can scan a text and say what it’s about, but analysis requires cognition, imagination, reasoning, inference and creativity – which is why my data science team includes people with backgrounds that range from math to music to forestry. (Yes, I have a colleague who can literally see the forest for the trees.)

The guiding principle for us is that we want to enable the discovery of knowledge. Insights are no longer as interesting as they once were. Big Data in action starts with the creation of a fact base to experiment with; ultimately, the goal is to be able to test hypotheses in real time, at the speed of real businesses. It’s not about knowing the right questions from the outset, but rather the process of finding your way to them.

That’s what I mean when I talk about Big Data.

JOSHUA SULLIVAN

Josh Sullivan is Senior Vice President of Strategic Innovation for Booz Allen Hamilton. Before that, he worked on government projects for private firms and was a US government engineer. He has a Master of Science degree in IT from Johns Hopkins University and a Ph.D. in Applied Computer Science from Northcentral University.

BOOZ ALLEN HAMILTON

Fortune 500 company Booz Allen Hamilton is one of the oldest management consulting firms in the world. It is also a provider of technology and engineering services to the US government in defense, intelligence and civil markets, and to corporations and not-for-profit organizations. The company focuses on improving efficiency, cybersecurity, healthcare and IT and delivers strategic innovation in areas such as analytics and data science. www.boozallen.com

Josh Sullivan spoke to **SARAH LUBMAN**, a Partner in Brunswick’s New York office.

ILLUSTRATION: DENIS CARRIER / SYNERGY ART