NASA Scientist Tells CIOs to Set Realistic Expectations for Big Data - Wall Street Journal (blog)

Big Data News Magazine February 4, 2015 Big Data News

SAN DIEGO— As analytics play a larger role in the corporate world, CIOs and other business leaders must establish realistic expectations about what they can hope to achieve with these tools, a NASA researcher advised. Working with Big Data requires an ability to accept uncertainty, and cope with the reality that even the most powerful analytic tools don't always yield satisfying results. "The peril is you spend time and don't get a thing out of it," said Dr. Amy Braverman, principal statistician at the Jet Propulsion Laboratory, which is associated with NASA. "The other peril is you spend a long time and get the wrong thing out of it."

Dr. Braverman develops new statistical methodologies for using datasets that are collected by NASA satellites and land stations on a range of missions and then stored in data centers around the country. Part of NASA's mission is gauging the health of the earth's atmosphere

through “thousands and thousands and thousands of variables.” Her models have real-world, political consequences. The goal is to “take an agnostic point of view, realizing that there are millions of people whose particular interest in that data I don’t know, but we need to make the data accessible to them nonetheless.”

Although (most) business data does not come from satellites or Mars Rovers scooping up samples of Martian landscape for analysis, Dr. Braverman’s warnings apply to corporate usage. Business data can be just as unwieldy. It can span a range of sources, and the way it is applied can have meaningful consequences.

Dr. Braverman also urged CIOs to avoid relying on familiar, off-the-shelf tools for analyzing data. They may need custom tools to solve their problems. “Don’t expect the existing software package you have is going to give you the tools to do what you need,” she said in a Big Data discussion at the WSJ CIO Network conference.

“Any time you compute data you collected, you are collecting sample statistics,” she said “It’s important to acknowledge there is uncertainty.”

For NASA’s carbon dioxide emissions project, there was uncertainty when it came to addressing the weakness of one of the carbon-dioxide-measuring satellites that NASA employs. The satellite runs along the polar orbit, crossing the equator and taking readings at 1:30 p.m. each day. It’s a time period associated with a cycle in which photosynthesis is peaking, and plants draw in a great deal of carbon dioxide. She said NASA had “just enough” land-based stations to develop some idea of what’s going on when the satellite is not taking readings.

Businesses likewise need to have a better understanding of what’s not being measured. The business equivalent of a land-based data collection station could be the focus group. “You could go into a great depth with a few people to understand the relationship between products at a far greater depth than opportunistic data collection of who buys what at grocery store,” Dr. Braverman said.

Probability models are extremely useful when it comes to addressing areas of uncertainty, she said. The key is approaching data from a quantifiable perspective.

Another important factor: Finding the right talent. “I think as a society we need to invest in the next generation of people of who are studying statistics and machine learning and data mining and all the tech that go into data science,” Dr. Braverman said.

Originally posted on http://ift.tt/1I3kPTR

Did you like this article? Share it with your friends!

Tweet 0 | +1 0 | Pin it

Leave a Reply

You must be logged in to post a comment.