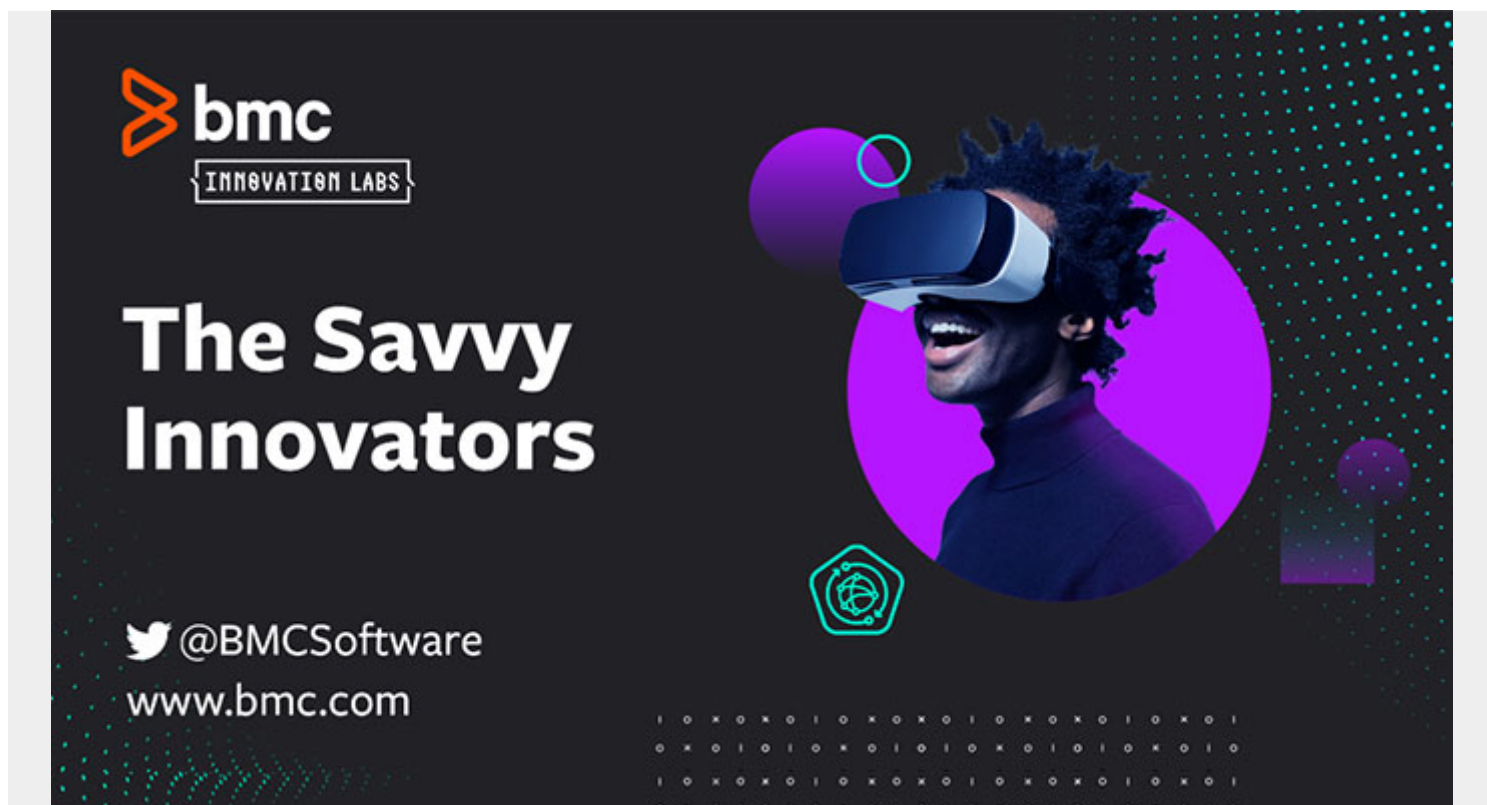


DELIVERING ON THE PROMISE OF YOUR DATA



Data is everywhere, in everything we do, but how do organizations turn that data into insights to become more data-driven and evolve to an [Autonomous Digital Enterprise](#)?

In our latest “Savvy Innovators” podcast, we were joined by thought leaders in the data space—[Jennifer Glenski](#), Director of Product Management in our BMC Innovation Labs; [Maria Glenski](#), Senior Research Scientist and Team Lead at Pacific Northwest National Laboratory; and [Phil Vincenzes](#), Chief Analytics Officer at IntelliDyne, to discuss “The True Value of Data.”

For the love of data

Phil began his work in the data field in the days after 9/11. “I built teams that pioneered the development of what we now call and know as open source exploitation. We couldn't wait to get to work in the morning,” he shares. “We were using data science, but we didn't know it at the time because that term didn't exist...to find the bad guys doing bad things. That was my defining moment for getting passionate about using data for national security purposes. And I haven't looked back since.”

Maria is part of a team that's taking on some of the world's greatest science and technology challenges. “I lead teams that are focused on research that models, characterizes, and explains complex systems and behaviors from humans in online and offline settings,” she says.

In her work, she looks at how information spreads and how users react to and consume different content, as well as how artificial intelligence (AI) fits into the mix. “The foundation of this work is the data. You can't go do any of it without the data that is observing the world around us, the

phenomena, how people interact, how humans interact, how AI and ML interact," she adds.

At BMC, Jennifer applies data science and analytics to building new products and finding out more about how BMC's customers are using data. "My passion for data really comes from being able to improve solutions and experiences for customers," she says. "Progress improvement, better design, all of that is where I get my joy of discovering and using data and where I see the value."

So much more than data

The value of data isn't in data for data's sake, but in all the things you can do with it. It's the entry point into a larger discussion. According to Phil, "Data, and more specifically, the study and applied use of data, has a real or 'potential impact' on every living thing and entity on the planet. When we're talking about the nitty gritty guts of computing, data is just a bunch of zeros and ones."

"It has no real meaning or value unless we trust it and then overlay some type of analysis or interpretation, translating it into something usable that gains knowledge. The value was held within the data. It's there, but it's waiting to be realized through. Our job as analysts and data scientists is to release that value that the data holds," he says.

Jennifer likens data to treasure that's waiting to be discovered. "That could be the application of that data perhaps an aggregate with additional data. And then your value extends past the initial application. So the value you get from using data isn't just the sum of all the pieces of data. The whole can be greater," she says.

She pointed to the example of a long-term medical research study she conducted. "There's a lot of historical cancer research. You might make a new finding or discovery that's valuable, but going on later, maybe a decade, you can look at the entire domain of that research. And some of those pieces can fill in the gaps of other studies, or you can see overarching trends that you wouldn't have gotten just from the initial collection and application of the data in the first place."

Maria looked at it another way. "There is value in data alone, but a lot of that value is what that data can support, what that data can help drive, develop, or provide insights on. But data alone is hard to have intrinsic value outside of that potential, especially if you're working with a lot of it," she says.

She explained that one of her projects has terabytes of data as part of an initiative to study math for artificial reasoning and science. "We're able to do really cool things with that. But those terabytes alone, if that was all that we were producing and it was just sitting on a shelf somewhere, how much value is there in that," she muses. "If we can have these large-scale data sets that are supporting continued development, continued advances in the field and in science, there's incredible value there."

The quality of data

Equally critical to how useful data becomes is its quality at the point of collection, something easier said than done when it's derived from human behaviors. "A lot of times, if you were working with large-scale data, especially human-generated data sources...humans are messy and AI and machine learning often don't love messy," Maria says.

"They want to have a nice, orderly, cleaned, consistent format to work with. And humans full grammatical sentences, and then we switch to looking on our phones, and...abbreviations, shorthand...all sorts of things. And it's harder for AI and ML to adapt to those kinds of different

switches. When you're working with AI and data science, that first step making sure that is AI-ready or in a format that your AI can take and run with."

"It's about having that data quality that you can trust," adds Jennifer. "And the data quality doesn't mean throwing everything out that looks out of line or abnormal. It means recognizing those change events or those anomalies. It means getting the data values within the appropriate range that you're expecting...before it populates throughout your organization and pollutes your data lakes or data warehouses and things like that, because it can be so hard to go back later and try and find all those pieces and scrub it and clean it up."

"You don't want to over- or under-engineer things, but you do need data pipelines and you need to automate them or orchestrate them, because if you're using that data to make important decisions or to improve customer experiences, those end results can have a significant impact on the success of your business."

Phil shared that he likes the phrase, "Crawl, walk, run," which means starting small and then demonstrating value before looking for the next big opportunity. He referenced a project with the United States Department of Justice (DOJ) where he was asked to help predict criminal and civil caseloads to appropriately staff almost 100 districts. He and his team analyzed the data in the DOJ's case management system and integrated it with publicly available open-source data. "We created a data model for staffing projections and it was very successful. But the interesting thing was that it unexpectedly predicted the opioid crisis, which was several years forthcoming," he says.

Data is the future

One of the tenets of the Autonomous Digital Enterprise is to become a [Data-Driven-Business](#) that captures, correlates, and monetizes data across the enterprise to yield high-value business cases with AI/ML while also optimizing and improving the processes of data extraction and analysis. So it makes perfect sense that people who know data best would be integral to that evolution.

"When we talk about citizen data scientists, to me, it just merely means providing some of the cool tools that make the analysis of information a lot easier. I think are always going to be completely in demand and are required, but we need to get more information and data value in the hands of the people that are making decisions," explains Maria.

She adds that the availability of new tools, Python programming resources, and computer science open sourcing of methods, data sets, and analysis notebooks is a great way to expand data science beyond the existing research community. As Phil says, "Now, everyone is clamoring to be the new thing. And the new kid on the block is really the data engineer."

Tune in [here](#) for the rest of the discussion and find out how our data thought leaders tied their concepts to songs by Justin Timberlake and Lizzo, and that iconic line from *JAWS*. And look for our upcoming research on the business value of data based on a survey of the current data practices of over 1,100 IT decision makers from around the world, coming in February 2023!