

COMMENTARY

Who is the Data Scientist?

By [Lindy Ryan](#)

DEC 19, 2013 5:00am ET

Data scientists have been called the sexiest new job in business intelligence. They are the stuff of dozens of articles, infographics and, on occasion, the beneficiary of a zinger or two about ponytails and hoodies.

In a dueling keynote at SPARK! Austin, Dr. Robin Bloor put on a pair of thick-rimmed glasses and morphed into a data scientist in front of my eyes. At Teradata PARTNERS in October, Neil Raden called them new versions of old quants. They've been firmly denied any of Jill Dyche's affections according to her blog post "Why I Won't Sleep with a Data Scientist." And, my personal favorite: industry writer Stephen Swoyer called them unicorns – fantastical silver-blooded beasts and impossible to catch.



The list of these emerging data scientist urban legends could go on. My point is this: Everyone who's anyone has something to say about a "data scientist." But for all their allegorical appeal, what – or who – *is* a data scientist? I talked to three, and a few other folks in the industry, and here is what I've found.

The Skillset of Data Scientist, Ph.D.

One often-disputed characteristic of the data scientist is their educational background and skillset, and there *are* some essentialities to being a good data scientist. They must be of an analytical and exploratory mindset; they must have a good understanding of how to do research; they should be possessing of statistical skills and be comfortable handling diverse data; they should be clear, effective communicators with the ability to interact across multiple business levels; and, finally, they must have a thorough understanding of their business context.

But do they *need* a Ph.D. – or, to put it another way, should our data unicorns be limited to Data Scientist, Ph.D. Dr. Alexander Borek of IBM – who has a Ph.D. in Engineering and Data Management himself – says that typically a Ph.D. student with statistical skills is a good candidate for a data scientist role – so long as he or she is willing to engage with the business context. Further, Boulder-based data scientist Dr. Nathan Halko – whose Ph.D. is in Applied Mathematics – says that math teaches you the ability to abstract problems and dive into data *without fear*. And, while a business background is important, it doesn't give a data scientist the skillset to execute a data problem. Nathan says that some of his contributions within his organization have been ideas he's hacked together over the weekend because he has the ability to *execute* – turning those ideas into *solutions* that can be delivered to others. Perhaps data people simply can more easily understand what the business needs than a businessperson can understand what the data is capable of?

Either way, competency in mathematics and statistics is unanimously important for a data scientist – perhaps *more* important than a business background. Yet, there's also a common sentiment that it's (typically) easier to interface with someone who has a business background, as opposed to a data scientist. And that's part of a business-education skillset: clear, effective communication delivered in a simple format that business executives expect – and that lacks some mysterious data jargon. Equally as important for the successful data scientist is the ability to translate and engage between both business and IT, and have a firm understanding of the business context in which they operate.

Why They Are/Aren't Unicorns

The two competencies are complementary, even if they are imbalanced. But there's a third skillset of these elusive data scientists that's a little more...intangible.

At this month's Big Analytics Roadshow in New York, a data science panel – comprised of data executives from Teradata, Comcast, Tableau and Radiant Advisors – had much to say regarding the perfect data scientist resume (though it was noted that the days of typing up a resume is “so twentieth century”). They also had quite a bit to say about what's *not* on the resume.

Teradata's SVP of Global Product Deployment and Strategy Tasso Argyros said that you don't need a degree that says Data Science to be a data scientist any more than you need an MBA. You need a foundational understanding of these concepts, yes, but more important, you need an eagerness to explore and discover within data. The characteristic that really seems to set the data scientist unicorn apart from the data user herd is their *personality*. A true data scientist possesses what John O'Brien of Radiant Advisors called a “suite of hidden skills,” including things like innovative thinking, the readiness to take risks and play with data, and a thirst to explore the unknown, and he looks to see how these skills are embedded within the blend of education and experience. Even in their own self-descriptions, data scientists seem to echo those same characteristics. Siva Yannapu of Blue Cross Blue Shield noted that data scientists are out-of-the-box thinkers; Nathan Halko described the data scientist as willing to have their hand – or hoof – in the metaphorical cookie jar (of data).

Being a data scientist isn't about checking off a list of qualifications and adding buzzwords to your resume. It's about *becoming* a data scientist – having the eagerness and hunger to dig deep inside data and find value. That's aspiration, and it's an intrinsic characteristic not taught in any program.

Sexy By Association

As far as I can tell, my little herd of unicorns doesn't find themselves all that sexy – or rare even. One said it's all the new *data* that's really sexy, not the guy in the glasses tinkering with it, thereby making the data scientist merely Sexy By Association. What they do think is that the role of a data scientist is a very interesting one that is intellectually challenging and can have a huge impact on the success of the business. Data scientists, according to Dr. Borek, are generating a quick and solid return on investment for most businesses and they will soon become a solid component of any larger business organization. Hence, it is a good future to bet on.

Sure it's sexy and maybe a little (or a lot) elusive, but we've got to start thinking of data science more broadly – not a particular set of technology or skills, but as those *people* who have a set of characteristics: curiosity, intelligence and the ability to communicate insights and assumptions. Those naturally inquisitive people already living and breathing the data within our businesses are every bit as much a data scientist, even if they don't have the fancy title.

One thing is certain: Data scientists come in many colors as the rainbows that their fantastical counterparts dance

upon. But data scientists, no matter how sexy or rare they are, aren't the only source of discovery within the organization, especially with the rapid increase of self-sufficient discovery tools allowing everyday business users to explore their own data. If you define data scientist community as a set of skills, you're missing out on a ton of people that already exist in your organization who can contribute a ton of value, too.

Author note: A Special Thank You to Nathan Halko of SpotRight, Dr. Alexander Borek of IBM, and Siva Yannapu of Blue Cross Blue Shield for contributing your thoughts and insights. You guys are unicorns in my book (except Nathan, who thinks unicorns have wings and is, therefore, a pegasus).

Lindy Ryan is Research Director, Data Discovery and Visualization, at Radiant Advisors. You may reach her at lindy.ryan@radiantadvisors.com or follow her on Twitter [@Lindy_Ryan](https://twitter.com/Lindy_Ryan)