EMC’s Bill Schmarzo and consultant Ben Woo weigh in on whether Big Data is revolutionary, evolutionary, or both.

by Terry Brown
EMC+
In a recent survey of CIOs conducted for EMC by CIO Magazine, almost half the respondents said they agreed with this statement: “Big Data analytics is an evolution but not a revolution in the area of data warehousing, databases, and big file systems.” The rest were about evenly divided among “game changer,” “pipedream,” “and not sure.” What’s up with that? Isn’t Big Data changing the world?

To add perspective to that decidedly equivocal survey result, EMC+ went to EMC’s Bill Schmarzo and Ben Woo of tech strategy firm Neuralytix. Schmarzo, a regular EMC Big Data blogger, works in EMC Consulting, and previously was Vice President of Advertiser Analytics at Yahoo. Before founding Neuralytix, Woo was Program Vice President of IDC’s Worldwide Storage Systems Research, where he launched the firm’s Big Data research effort.

We asked Schmarzo and Woo what they thought of that survey result, and neither was surprised.

Schmarzo: In fact, I would have thought the differentiation would be even been higher. I think the problem our industry faces is that we focus in on the Facebooks, Yahoos, Zyngas, Googles, and Twitters out there, who built this technology for themselves because existing platforms didn’t work well for them. But those stories don’t resonate with 99% of the marketplace.

The companies I talk to have almost no interest at all in what Facebook or Yahoo’s been doing. It almost has no relevance to them. Whether it’s insurance policies, or tractors, or toys, or shoes, whatever it is, the stories you hear are about companies whose only business is data. Too many customers get the impression they can get to the moon by climbing to the top of a tree.

EMC+: So is Big Data revolutionary or evolutionary?

Woo: Big data is not a new technology. It’s actually just an improvement on the old. So the people who answered the survey are not entirely wrong in having that position. Nothing about big data and its processes are anything revolutionary from that perspective. What is revolutionary and game-changing is that for the first time we have the necessary compute power, the necessary memory and network, the storage, and most important, the necessary software to actually consider the entire set of data.

Schmarzo: It’s evolutionary in the technology, but it’s game-changing how it’s applied to the business. That’s what makes this so interesting, that we’re not talking game-changing from a technology perspective. That is, leverage your BI and data warehouse assets to get more out of them. We’re talking game-changing in how you deploy it at the point of customer engagement.

The challenge is that companies that don’t do this will be out of business. Big Data is an evolutionary game-changer, where companies are figuring out, how do I bring data into my product to make it more effective, more productive, more relevant to the user?
EMC+: So we have a semantic problem here? Because there is something revolutionary going on, somewhere. Is it a revolution in attitude and vision more than anything else?

Woo: I think that certainly is a start. One of the challenges we have is the fact that we haven’t been in a position to do this before, as we didn’t think it was possible. We always thought it was too expensive. And in many situations, that was absolutely true. So the difference here is, it’s no longer expensive. We don’t need to go out and buy a multi-million dollar specialized system. And we can do this at close to real time. You know, for many, many years, we used to put all this data into a data warehouse. Then we would do a separate process to mine that data and look at it and make assumptions about it. With big data, we can do a lot of things in parallel.

Schmarzo: You have to see the data you have as an asset, do some level of intelligent transformation, which is heavily analytical and data-centric, and turn it into something that customers want. That’s the start. Maybe it’s turning visitors into audiences or turning sites and properties into inventory. That is what data monetization is all about. Do some creative analysis with your data to create something that your users want to buy.

Whether it’s on the cloud or cheap scale-out architectures on commodity microprocessors, I can be a regional retailer with maybe 100 stores and I can build out a big data platform using x86 processors, Hadoop, and an MPP Postgres architecture. It doesn’t have to cost me very much money.

EMC+: So you can spend a little money, maybe an evolutionary amount of money, to make a lot?

Woo: We’ve been looking at cost efficiencies since day one of commercial computing. IT spend, for most organizations, is relatively minimal. We’re talking about low single digits spend on IT as a percent of revenue. So let’s just call it 2%, which is a number that’s relatively well accepted.

If I save your company 10% on the technology, I’ve moved from 2% of revenue to 1.8% of revenue. That’s nice, it’s important, don’t get me wrong. But if I can take technology as an investment and find or help generate new lines of revenue, new products, or even more important, create some competitive advantage as a result of it, now we’re talking about moving tens of percent on the top line.

And when you start looking at that, when you start moving the needle on revenue and profitability, just by moving the revenue up, you’ve naturally moved the percent of technology spend down.
EMC+: So basically, it's just understand your business, and fantasize about the ways you can make your customers happier. And out of that is very likely to come a use case for big data.

Schmarzo: Exactly. An auto manufacturer, a telecom, a retailer. Look for opportunities to leverage this data to provide a much better customer, or user, or shopper, or driver experience. That, to me, is game-changing. It’s evolutionary in a sense that I’m not asking people to change their business models. We’re asking them to take advantage of the data they have already around them to provide a much better experience for their consumers, for their employees, for their partners, for their suppliers, whoever it might be within their ecosystem.

What it costs me is having the right kind of creative people who can take what assets I have and look at what my customers are trying to buy, and make that transformation. It may be taking advantage of the data I have, or maybe even to instrument further to get more data. It probably relies heavily on doing some analytics, so I can take that data and turn it into something that’s useful.

Because none of our customers want to buy data. They want to buy insights. They want to buy things that make them more productive, or makes data easier to use, or creates a simpler experience.

EMC+: What is the greatest barrier to entry for a mid-sized company?

Schmarzo: Mindset and developing a Big Data skill set. In terms of mindset, it’s sitting down and really thinking through what your customers want. I think that's the biggest challenge, because the technology is out there. As for the skill set, if you have to rely on Hadoop to do some of this stuff, which today you have to in many cases, it means you have to find or develop some level of Hadoop skills.

In fact, this is, I think, a problem that all organizations face, which is, how do I re-task the skill sets I have? I've got all these BI and data warehouse people. They understand how my customers think. They understand how to interface with them. But they don’t know how to use some of these other new tools that allow them to glean new insights out of this data.

So there’s a mindset challenge, and I think there's a skills re-tasking challenge of how do I move people from where they are today to give them some new skills. We’re not asking them to scale Mount Everest. I think people started off by coding in COBOL. They probably started in Fortran, to COBOL, to C++, to Java, and now we’re saying you’re going to learn how to write Perl and MapReduce. It’s just coding.

Woo: Big data is not a technology decision. Big data is a business decision, a decision to say, we believe in our customers and we have information about them we can use. Every enterprise has information about their own customers. The ability to mine that data, the ability to extract and abstract new ways of keeping that customer—that is why you engage in big data. Generally IT is not the first to be initiating this stuff. This is initiated from marketing or from sales or from the CFO or the CEO.
And that's where companies need to recognize, early in the conversation, that technology is not just a spend. It is an opportunity to make money, and for the executives or the business lines execs to ask themselves what do we have already that can make us more money or have a better understanding of our customer? So those are the first questions to ask.

**EMC+:** So you’ve got your mindset right. You’ve got some bright ideas what you might do with your data. Can just about anybody start small and accomplish something and build from there?

**Schmarzo:** Well actually, the key thing is that most organizations already have the building blocks to have a big data game-changing experience. Let me give you an example, going back to retail. Most companies have captured and stored, probably on tape now, all their customer loyalty data, transaction data, right? They’ve got all this data.

Now historically, they’ve aggregated it for reporting purposes, and so they take in all the detailed point-of-sale data that came off the cash registers, and they aggregated it to show, what are my sales by store, by product category, by product, etc. And they lost the nuances in the data, because they’ve aggregated this data. But they still have that detailed data out there, but they’re working off a data warehouse platform that can’t handle that detailed data.

Well, guess what, folks? The platforms that can handle that detailed data? They’re out there. We happen to have, in Greenplum, one of the better ones. Not only because we’ve got the same sort of scale-out capabilities, but because of the software-only solution, which really allows companies to get in really cheap and grow incrementally, without having to take that big giant hurdle of buying some Superdome from HP.

**Woo:** Yes you can start small, but the thing is to start. I guarantee you that your competitors are doing something with their data. The most valuable thing any organization has are the records it holds, the transactions, and the digital data that they have. It’s a goldmine waiting to be mined. If you can understand your own customers better, you’ll have much lot better chance of understanding like-minded customers and prospective customers.

**EMC+:** You’re going to need people to implement the program. Do you buy those people or do you grow them?

**Schmarzo:** It depends on what you’re doing and who you have on your team—some will have to do both. But I think to start you need to think about the assets you have already. Can you grow these people? Because they’re the people who understand how the organization works. If you bring in people from the outside, you bring in people who have different business model exposures. Do you need that? You can always teach technology. You don’t necessarily need to start with a whole new team.
Woo: Let me give you an example. A major multi-billion dollar manufacturer looked at their own data in terms of revenue and customer satisfaction, etc. They wanted to find deeper metrics about measuring themselves. But as they went deeper, they found that they actually had more data on field service and field support of their products than they realized. Ultimately they were able to do predictive analysis on failure rates to the point that they could predict with a fairly high degree of certainty the possible failures that are likely to happen as a result of wear and tear and other things at their customer sites. Now they could be preemptive and proactive about maintenance. They grew that insight from inside, with their own people, using data they already had.

EMC+: How do you get the momentum going? Who starts the ball rolling?

Schmarzo: This stuff is only successful if it's used to solve some very interesting problems, or is used to enhance the value of an organization's value chain. So you've got to have an executive sponsor, and you've got to have somebody who understands that it's not about technology, it's about business. The technology's only an enabler. I guess that gets back to the point. Is it a game-changer or is it evolution? The answer is it's both. And the companies that'll be successful, the vendors who'll be successful in this space, are the ones who are able to convince IT that it's an evolutionary move that's going to build off the assets you already built, but who can also convince the business executives it's a game-changer in how they reform their value chain, and how they integrate and service their customers. That's a tough line to walk. But once you do, you can build a more competitive business.