Big data: Why it matters to the midmarket

Big data. The concept has been ubiquitous lately. It seems everyone has been talking about the use of big data to advance scientific research, perform complex number crunching, and predict presidential elections. There's little doubt that the chatter around big data has reached a fever pitch.

But one place where big data hasn't seemed to make much noise is among small and midsize companies. Perhaps it is widely assumed that big data, which imbues a sense of grandiosity, is only for those large enterprises with enormous amounts of data and the dedicated IT staff to tackle it. Smaller companies couldn't be blamed for thinking that big data, as the name suggests, is only for the big boys.

In fact, big data – and the analytics that make sense of it – applies to companies of all sizes. Rebecca Shockley, Global Research Leader for Business Analytics at the IBM Institute for Business Value, is one of the authors of the 2012 report, "Analytics: The real world use of big data." Shockley argues that one of the misconceptions of big data is that it refers only to massive quantities of data. She believes that the definition of big data depends on who is doing the talking.

"Big data is quite simply data that cannot be managed or analyzed by traditional technologies," says Shockley. "So what is considered big data for one company may be different for another company. 'Big' doesn't have to be really that big; it's just bigger than what you're used to dealing with."

Today, the ability to capture, store and analyze larger quantities of data is well within reach for midsize companies. As the world becomes more digital and more data becomes easily available, midsize companies in particular are poised to make use of all that additional information. Big data is no longer just a curiosity for midsize firms, it's an important new resource that can be transformative.

Managing uncertainty around big data
"A lot of the business intelligence tools that have come before this latest era of transformation, while very good and still extremely relevant to business operations, were never designed to handle what we call the characteristics of big data," says Shockley.

> ## "Big data is quite simply data that cannot be managed or analyzed by traditional technologies."

There has been little argument that big data is the confluence of three characteristics, or what have come to be called the 3 V's: volume, variety and velocity. To translate, big data is not only big; it also comes from a wider variety of sources including unstructured data like e-mails and videos, and comes in at a rate faster than traditional tools can handle.

But Shockley would argue that big data is as much about a fourth V, veracity, which can be often overlooked but just as critical. Some data, such as tweets, sensor-generated data and weather reports, can be inaccurate and brings with it an amount of uncertainty. "There's a certain sense of unreliability to some types of data," says Shockley. "Under normal circumstances, that unreliability, as a single data point in more traditional BI tools, could be kind of discounted or just treated like it was a certainty. But when you're dealing with big data, that unreliability begins to compound upon itself, and so we feel you really need to focus on how to manage the uncertainty."

This uncertainty presents a unique challenge to traditional analytic tools. Unlike other analytic initiatives, big data can require highly sophisticated mathematics that can manage any inaccuracy inherent in the data. For midsize companies that are looking to tackle big data initiatives, there are new technologies, namely analytic accelerators, which can easily handle complex analysis.

Shockley explains, "Analytic accelerators take the inherent complexity of analyzing big data and put it behind the screen so that the end user is just focused on inputting the facts, while the algorithm does all the heavy lifting for them. For small and midsize businesses with limited IT staff and resources, these analytic accelerators give them a way to overcome certain skills gaps."

Getting everything in order

Shockley recommends that the first thing companies should do before embarking on big data is to get their data in order. Many midsize companies have the capability to view different aspects of their business, but they fail to understand the relationship between them. "In a lot of organizations, they can look at products, they can look at customers, and they can look at sales, but they can't tie those things together. Being able to look at those things in an integrated fashion is really the foundation you need."

And big data is, above all, about volume. Companies should consider ensuring that they have the infrastructure necessary to handle the requirements associated with larger quantities of data. Shockley explains, "The first step is scalability. We found that volume is the first priority that most organizations are focused on, so make sure that you have a scalable storage infrastructure. And the next step on infrastructure tends to be the high-capacity warehouse which really means they put in more processing power so that they can handle processing larger volumes of data."

Once everything is in place, Shockley recommends that midsize companies focus on extracting insight from their internal data, rather than data from outside the four walls. "There is so much

untapped information just sitting in the data that you already have. In the beginning, organizations really don't need to try to tackle understanding the new tools, understanding the new analysis and all of this external data that is largely unstructured all at once. If you can look at your internal data across lines and process it at a scale that you haven't been able to do before, you can open up innumerable insights while getting your feet wet with the new tools and analysis techniques."

Start small

On a smarter planet, the use of big data will inform and empower companies to serve their customers and run their business in ways currently unimaginable. Shockley advises that as a first step, companies need to start thinking about the first question that they want big data to answer. "Start with the business problem that you're having. Are you trying to get new customers or grow the customers that you have? Work through what exactly it is that you're trying to do, and then go back and figure out what data you have that might answer that question, what data you need to answer that question, and how your data needs to be put together to make that work." Once you've tackled the process once, Shockley argues, you have a clear blueprint of how to successfully answer other questions critical to your business.

Because the potential for big data is often spoken about in such grand, earth-shattering terms, it feels vaguely oversimplified for companies' to start with just one simple question. But every important journey starts with one first step. "Data doesn't really create any value until you put it into action," says Shockley. "So start by understanding what action you want to take."

Learn more by going to:
ibm.com/businesscenter/forwardview