***Big Risks.** Vast stores of information can provide organizations endless insight on their business. Managing and safeguarding all that data is another story.

Stories are emerging from the campaign trail of last year's U.S. presidential elections about number-crunchers who set aside traditional polling and instead used vast amounts of information—data from such diverse sources as home ownership records, voting histories, and even logs of magazine subscriptions—to predict what people thought about the candidates, whether those perceptions could be changed, and what it would take to do so. They didn't use any new bits of data; after all, people can only be characterized by the data they generate in a limited number of ways. Rather, the analysts in the 2012 election used far more of those bits—far, far more. They used "big data" to figure out how the election would go.

And they were right, when the traditional pollsters were not. Indeed, as the *Los Angeles Times* reported, math wizards such as *The New York Times'* Nate Silver "champion[ed] statistical methods and advanced computing power" to paint a more accurate picture of how the electorate would respond to the candidates' appeals. What those experts have in common is their reliance on volumes and volumes of data—a reliance that led one columnist before the election to mock the Obama campaign's focus on data analysis as "politics as done by Martians." Now commentators remark that big data likely has changed forever the way elections are run.

For internal auditors, a key takeaway from the election is that the rise of big data is real; it's not a buzzword anymore. Like political campaigns, many businesses are looking to it to tell them more about their customers, their buying habits, and the likelihood they'll behave in certain ways under certain circumstances. Indeed, consulting firm McKinsey & Co. recently reported that some two-thirds of C-suite executives surveyed consider leveraging big data to be a top strategic priority, as they

Russell A. Jackson

BIG DATA*

seek to gain customer insights, effect more accurate budgeting and better performance management, and develop new products. Trouble is, McKinsey reports, huge gaps exist between what organizations hope to accomplish and what they're able to, given their existing IT infrastructures and expertise. That opens the door to all kinds of new risks.

DATA, IT'S REALLY BIG

It would be helpful to start with a definition of *big data*, but the concept is still too new to be defined consistently. Cody Sudmeier, a principal at Agility Solutions, an analytics and revenue assurance company based in Denver, suggests that internal auditors think of big data as not just a single set of data, "but the way data grow when you realize the ways you might be able to connect different sets of data together to create even more sets of information." For auditors, the important thing is to "make sure you're clear on the definition being used by your organization," notes Jeff Spivey, president of Security Risk Management, a Charlotte, N.C.-based consulting firm.

For their part, the consulting companies tend to define big data in terms of "bigger" rather than "different." McKinsey says the term refers to gigantic, complex data sets that are beyond the analytical capabilities of the IT tools available to most companies. Indeed, a report from consulting firm Protiviti Inc. notes that these data sets typically overwhelm traditional analysis tools such as spreadsheets and relational databases. Those data include information generated by customer transactions, visits to company websites, and company-related activity on social media.

"Big data represents the ever-expanding collection of data sets that sheer size, variety, and speed of generation make difficult to manage and harness information from," Matt McGivern, an Atlanta-based managing director of business intelligence and data governance at Protiviti, explains. "The sources of these electronic data are varied, being as simple as transactions in a banking system or as complex as nonstructured content such as Facebook pictures and videos." Because big data is characterized more by size than by content, the challenges organizations face when dealing with it have changed dramatically over the years, from needing to address the overall costs of physical storage to, now, needing to deal with the explosion of stored content that results from the abundance of cheap storage, he adds.

For internal auditors, a big issue with big data is determining what their organization does with it.

Eric Ginsburg, internal audit engagement principal at Comcast Corp., a digital media and home security company based in Philadelphia, doesn't worry about a precise definition of big data. "When I think about it, I'm thinking about business intelligence, which goes beyond financial reporting systems and business intelligence tools that just deliver reporting from your general ledger," he says. "It's intelligence about your customers and trending the various aspects of your business. It's bringing social media data and correlating it to some of the trends in your business or in the market." That's a world away from what internal audit normally does, he adds. "Big data is all of those really big, really complicated data sources that financial and accounting professionals have not encountered before, but that are becoming increasingly important." Ginsburg concedes that auditors haven't been thinking about such information up to now, but he emphasizes, "we shouldn't be closing any doors."

One thing most experts agree on is big data is not simply data mining on a greater scale. McGivern points out that data mining typically is conducted on more structured data from a limited set of sources, such as transaction logs from a single server or application. "Due to the larger number and variety of sources for big data, most traditional data mining tools and techniques are inadequate," he says. And James Riley, senior data analyst at VF Internal Audit, in Greensboro, N.C., considers data mining a component of big data. "Think of big data as a survey or prospecting rather than as mining," he explains. "Sometimes technology can offer the ability to drill down or zoom in and instantly open your mine. Data mining will typically have an objective and a predictable outcome that can be tested."

A UNIVERSE OF RISK

For internal auditors, a bigger issue with big data is determining what their organization does with it. There are several key risk areas that internal audit departments should expect to address when their organization takes the big data plunge.

Privacy and Security As Spivey points out, "one concern is big data produces intelligence you didn't have before." As such, internal auditors will need to educate themselves—so they can educate management and the audit committee—on whether and how the

By 2015, **consumers** will create **68%** of the world's unstructured **data** (e.g., from search results and social networking), Booz & Co. predicts – growing 15 times the rate of structured data.

new data are covered under both new and existing privacy laws.

Sudmeier's colleague Sami Ibrahim points out that "high-quality, well-organized, and reliable information is valuable. The problem is it's also valuable outside the organization—whether it be customer lists, pricing, order details, financial reporting, research and development concepts, or employee salaries. When compromised, those information sets have resulted in huge fines and immeasurable impact on reputation." While it would be nice to have the foresight to see all potential risks, Ibrahim says a more reasonable approach is for internal audit simply to keep up with the events that have already been reported in the media. "These actual events provide a great framework for identifying the potential sources of risk," he says, "and, more importantly, a tangible means of communicating the potential impact to management."

Appropriate Data Destruction Any company that uses data—big or not, digital or not—should already have policies in place for getting rid of what it doesn't need anymore. Just as with paper documents, Ibrahim notes, a comprehensive data management policy is the foundation of managing the risks of inappropriate data disposal, and internal auditors should ensure that their organization has assigned ownership of the disposal policies to an appropriate department or individual—often the chief legal officer—and that the policies cover all necessary sets of information.

In addition, internal audit can help ensure that the policies are communicated regularly and that management tests whether the policies are followed. "This can be a tricky area," Ibrahim says, "as data retention requirements can be very specific to

BEYOND DATA ANALYTICS

The advent of the big data era doesn't just affect the risks that internal audit examines. It also offers internal audit departments new tools to perform familiar functions. Agility Solutions' Cody Sudmeier notes that, when used correctly, big data can give internal audit the ability to help the organization by:

- **Measuring potential risk exposure.** Rather than relying on sample-based techniques, internal audit can provide a more reliable measurement of the extent of an issue by using data analytic tools.
- **Sourcing the risk and identifying corrective actions.** Often, large sets of information can reveal patterns that would not be observed on a more limited basis.
- **Helping evaluate the costs and benefits of control improvements.** With a more reliable measurement of a risk, management has better information for evaluating the costs and benefits of control enhancements.
- **Monitoring.** Using analytics, internal audit can monitor the effectiveness of controls on a complete, yet cost-effective basis.

Indeed, as Protiviti's Matt McGivern points out, "big data offers internal audit a unique way to get a more holistic view of key events that they may be reviewing." For example, often internal audit will only have a portion of the information regarding a fraud or adverse event, such as an application transaction log. "Imagine the benefit to internal audit if it could see—all within the same reporting—the Web server events, correlated with the firewall events, operating system events, and application server events," he says. "That combined picture offers a much richer view of the event."

How can internal audit make that happen? Ginsburg says the first step is to make sure it's a stakeholder in the development of data warehouses and analytical systems. "Be a part of the user group," he advises. "Be in on the conversation so you don't have a retrospective view of what was done well and what wasn't." Also, when possible, internal audit should develop its own team of specialists with data analytics skills who can focus on the department's needs and help it establish a viewpoint for how the data should be used. "The way internal audit uses data is always a little different from the rest of the company," he observes.

"When compromised, information sets have resulted in huge fines and immeasurable impact on reputation."

Sami Ibrahim

> "Auditors should work with management to determine the strategic purpose for capturing these large data sets."
>
> Matt McGivern

the types of information covered, the industry, and other criteria." Also, organizations may be tempted to keep all their data, as storage costs get cheaper. "However, keeping unneeded data, or not having it stored logically, can increase the risks and costs of legal discovery," he emphasizes.

Lack of Internal Expertise or Leadership "Most organizations probably don't have the experience they need to leverage the financial and operational data that's already available, let alone big data," Ginsburg notes. The risk of that, McGivern points out, is "organizations may inadvertently be capturing information that they should not be storing, or they may be missing key indicators of business risk to which management should have been alerted. Both of these cases introduce additional risks to the organization and should be addressed by internal audit." Specifically, he says, data shouldn't be captured without some basic information management concepts in place, such as data classification and data profiling schemes.

Additionally, internal audit should review the sources for data capture to ensure that violations of customer privacy are not occurring and should educate management on all of the risks involved in capturing and storing big data. "Auditors also should work with management to determine the overall strategic purpose behind capturing these large data sets," McGivern stresses, "and ensure that the data are being stored in line with regulatory retention and privacy rules."

FILLING THE GAPS

Its sheer volume can make big data very intimidating, even at the conceptual level, but at the end of the day, it's just another business risk. Sudmeier points out that different organizations—and internal audit functions—adapt to new technologies at different paces. "That being said, most shops see the potential value in being able to tap into big data in their organizations to help stay current with business activities," he says. "Internal auditors can play a major role in testing the quality of data used in operational, compliance, or decision-making contexts."

When it comes to big data projects, internal audit needs to have a seat at the table and ask hard questions about their risks and rewards, McGivern advises. "Internal audit must be active in the up-front profiling and data classification for all data sets contained within any big data projects, which will help the organization determine the risks faced through loss or theft of the information," he says. Additionally, auditors should assess and advise the organization on technical safeguards and controls that need to be in place for big data projects, as well as review the overall strategies that the organization is considering before launching such initiatives. Sources for the big data efforts also should be outlined, and internal audit should be made aware of any contractual restrictions for data pulled from those sources, he adds.

As in all cases, Spivey says when big data projects are on the drawing board, internal audit's responsibility is ensuring that risks are understood and that there's a process in place to understand and manage them. "Every time there's a new 'thing,' whether it's big data or cloud computing, there needs to be a complete understanding of how to map what the organization is doing inside the new environment," he says. "We may find out there are gaps between what we did in the past and what we need to do in the future. Internal audit makes sure those gaps are being filled."

RUSSELL A. JACKSON *is a freelance writer based in West Hollywood, Calif.*