



*For the Complete Technology & Database Professional*

# **BIG DATA, BIG CHALLENGES, BIG OPPORTUNITIES:**

## **2012 IOUG BIG DATA STRATEGIES SURVEY**

---

By Joseph McKendrick, Research Analyst  
Produced by Unisphere Research, a Division of Information Today, Inc.  
September 2012

Sponsored by **ORACLE®**

Produced by



Thomas J. Wilson, President

---

## TABLE OF CONTENTS

|  |           |
|--|-----------|
| <i>Executive Summary</i> .....                       | <b>3</b>  |
| <i>Data Environment</i> .....                        | <b>4</b>  |
| <i>Business Drivers</i> .....                        | <b>12</b> |
| <i>Big Data Management Tools and Platforms</i> ..... | <b>19</b> |
| <i>IOUG Recommends</i> .....                         | <b>29</b> |
| <i>Demographics</i> .....                            | <b>30</b> |

## EXECUTIVE SUMMARY

As more data becomes available from an abundance of sources both within and outside, organizations are seeking to use those abundant resources to increase innovation, retain customers, and increase operational efficiency. At the same time, organizations are challenged by their end users, who are demanding greater capability and integration to mine and analyze burgeoning new sources of information.

Big Data provides opportunities for business users to ask questions they never were able to ask before. How can a financial organization find better ways to detect fraud? How can an insurance company gain a deeper insight into its customers to see who may be the least economical to insure? How does a software company find its most at-risk customers—those who are about to deploy a competitive product? They need to integrate Big Data techniques with their current enterprise data to gain that competitive advantage.

The opportunities and challenges presented by Big Data are addressed in a new survey of 298 data managers and professionals who are part of the Independent Oracle Users Group. The survey was underwritten by Oracle Corporation and conducted by Unisphere Research, a division of Information Today, Inc. Survey respondents hold a variety of job roles and represent a wide range of organization types and sizes and industry verticals. The largest segment (36%) of respondents hold the title of database administrator, followed by that of director or manager. Close to one-third work for very large organizations with more than 10,000 employees. By industry sector, the majority of respondents come from IT service providers, financial services, education and government agencies. (See Figures 36–38 at the end of this report for more detailed demographic information on job titles, company sizes, and industry groups.)

While definitions of “Big Data” vary, it is generally seen as increased volume, variety, velocity, and value. For purposes of this survey, we looked at two of the key differentiators of Big Data versus traditional data stores—volume and variety.

**Key highlights and findings from the survey, which explores Big Data issues and solutions, include the following:**

- More than one out of 10 data managers now have in excess of a petabyte of data within their organizations, and a majority of respondents report their levels of unstructured data are growing. Fewer than one out of five feel their IT infrastructure will be ready to handle this incoming surge of data. Protecting data overall is important, but unstructured data gets low priority at this time.
- While many organizations are still struggling to understand the business value of Big Data, more than 50% consider it to be extremely or very important to their business. The greatest opportunities Big Data offers are in competing more effectively (40%) and growing business revenue streams (32%). Most Big Data initiatives currently come out of the IT department, but the first business applications are being seen in marketing and sales—an area typically already comfortable with data analytics.
- Hadoop—often seen as the cornerstone of many Big Data efforts—has yet to catch on within most Oracle environments. This suggests that many organizations are just starting to address Big Data storage requirements. Adoption of Hadoop, in fact, is expected to double over the coming year.

The challenges—and enhanced expectations that Big Data brings—were expressed by one respondent, a project manager with a global financial services company. “The growth in data has been incredible,” she says. “Our executives expect the data faster, sliced and diced multiple ways, and summarized into executive summaries, dashboards, PowerPoint presentations, and white papers. Speed and accuracy are of the essence in winning new business and maintaining current customers.”

On the following pages are the results of this latest examination into the emerging world of Big Data.

## DATA ENVIRONMENT

**More than one out of 10 data managers now have in excess of a petabyte of data within their organizations, and a majority of respondents report their levels of unstructured data are growing. Fewer than one out of five feel their IT infrastructure will be ready to handle this incoming surge of data. Protecting data overall is important, but unstructured data gets low priority at this time.**

There are multiple ways to measure Big Data—which can be based on volume, variety, velocity, and value. For the purposes of this survey, we looked at two of the key differentiators of Big Data versus traditional data stores—volume and variety.

In terms of volume, the survey finds considerable amounts of data now being supported within today's Oracle enterprises. For instance, 11% of respondents have data stores within their enterprises that exceed the one-petabyte mark. Another 20% report they are managing data in the hundreds-of-terabytes range. Overall, 42% can be considered large data shops, supporting more than 50TB. (See Figure 1.)

Of course, these levels vary greatly by company size. For example, 28% of the largest organizations in the survey (with more than 10,000 employees) report having more than a petabyte's worth of information in their shops, compared to only 1% of the smallest firms with fewer than 1,000 employees. (See Figure 2.)

Another measure of Big Data is variety, as seen in the degree of unstructured data (web logs, social media data, sensor data, documents, imagery, audio) coursing through enterprise systems. At this time, close to one-fifth of the enterprises surveyed say a significant percentage of their data (25% or more) is unstructured. (See Figure 3.) Even more telling, close to two-thirds of respondents indicate they expect the amount of unstructured data in their organizations to increase over the next three years, either significantly or moderately. (See Figure 4.) The leading data types that comprise the growing Big Data stores include transactional data, email files, and office documents, the survey finds. (See Figure 5.)

While many respondents are bracing for the Big Data deluge—if they aren't already in the thick of it—current systems, as they are configured, may not be ready for the onslaught. Most data managers in the survey, 65%, feel that for the most part, their current IT infrastructure and database systems are adequate for managing all their data—at this time. However, when they

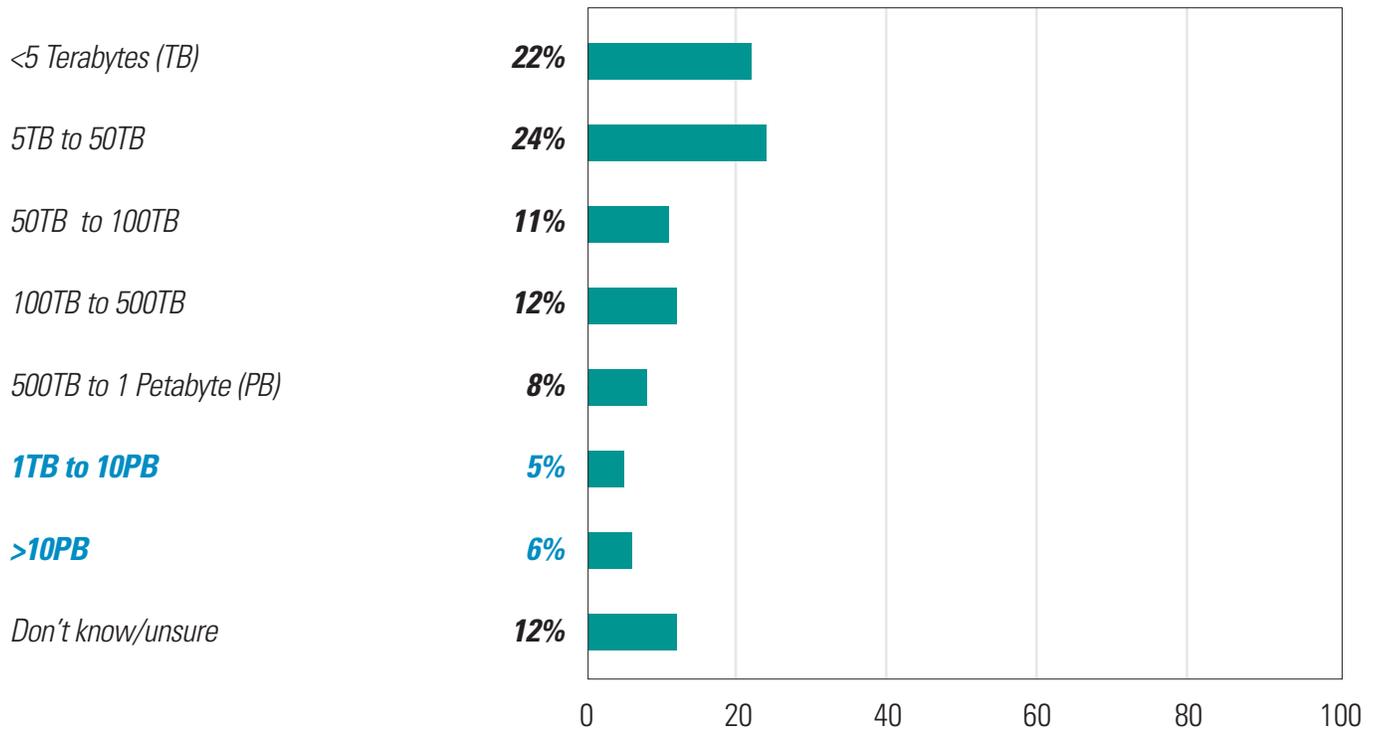
consider their systems' adequacy to handle data management requirements three years from now, the percentage expressing such confidence drops to 54%. In addition, this confidence is lukewarm at best—only 19% indicate they feel completely assured that their IT and data infrastructure will be up to the task. (See Figure 6.)

The sense of inadequacy is even more intense among the Big Data organizations identified in this survey. Only 11% of respondents in Big Data sites—defined as those reporting more than 500TB and more than 25% unstructured data stores—are confident their systems will be up to the task in three years. (See Figure 7.)

While managing all this data is one thing, another key consideration with these growing volumes and the variety of data is its criticality to the business. Keeping data highly available and secure is an ongoing challenge for data managers. How much data presented to respondents' infrastructure and database systems can be lost without repercussions to the organization? For example, can the business afford the loss of a store of unstructured data, such as graphics files, such as videos or web logs? More than one-third, 37%, indicate that absolutely none of their data can be lost. In total, 65% indicate they can't sustain data losses exceeding 5% of their total information assets. (See Figure 8.)

While there is a drive to guard almost all data against loss, levels of protection vary significantly by data type—as indicated by half the respondents. (See Figure 9.) And, accordingly, unstructured data receives low priority on the data protection spectrum. A majority of respondents, 77%, consider transactional data to be most important. Just under half make every effort to protect the integrity of office documents stored on their premises, and about 45% consider their email to be too important to lose. One out of five say it is important to protect device-generated or locational data, but only one out of 10 worry about web logs and audio or video files. (See Figure 10.)

## Figure 1: Total Amount of Data Managed Today



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

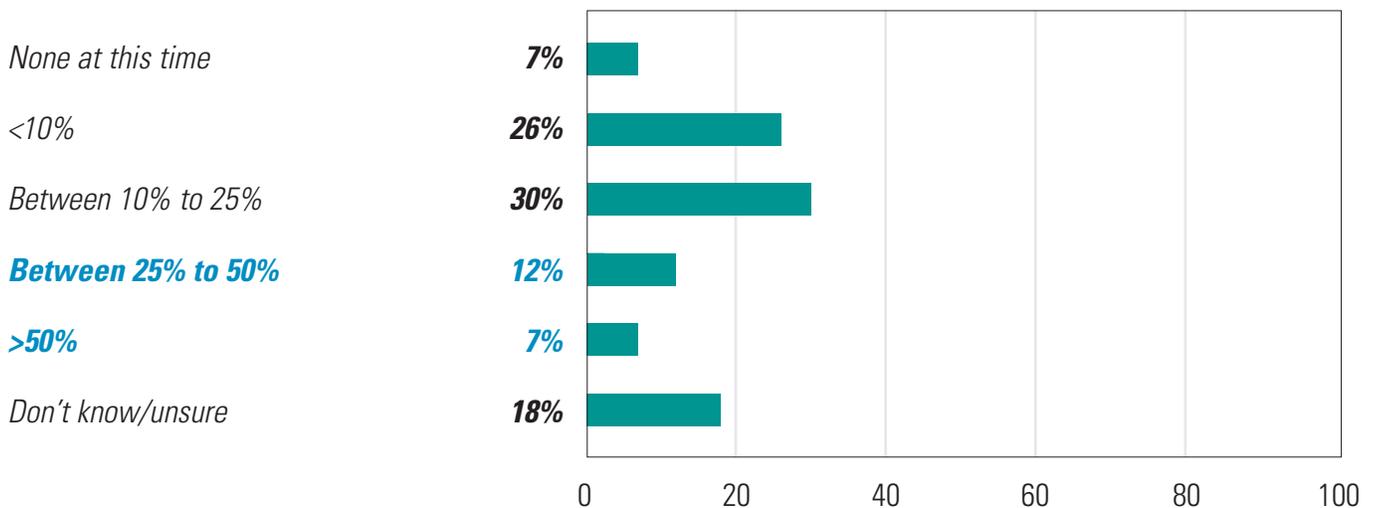
Data collection and analysis performed with SurveyMethods.

## Figure 2: Total Amount of Data—By Company Size

|                          | <1,000 employees | 1,000–10,000 employees | >10,000 employees |
|--------------------------|------------------|------------------------|-------------------|
| <5 Terabytes (TB)        | 37%              | 11%                    | 12%               |
| 5TB to 100TB             | 41%              | 45%                    | 24%               |
| 100TB to 500TB           | 10%              | 18%                    | 9%                |
| 500TB to 1 Petabyte (PB) | 2%               | 7%                     | 14%               |
| >1PB                     | 1%               | 8%                     | 28%               |

("Don't know/unsure" responses not included, so totals do not equal 100%.)

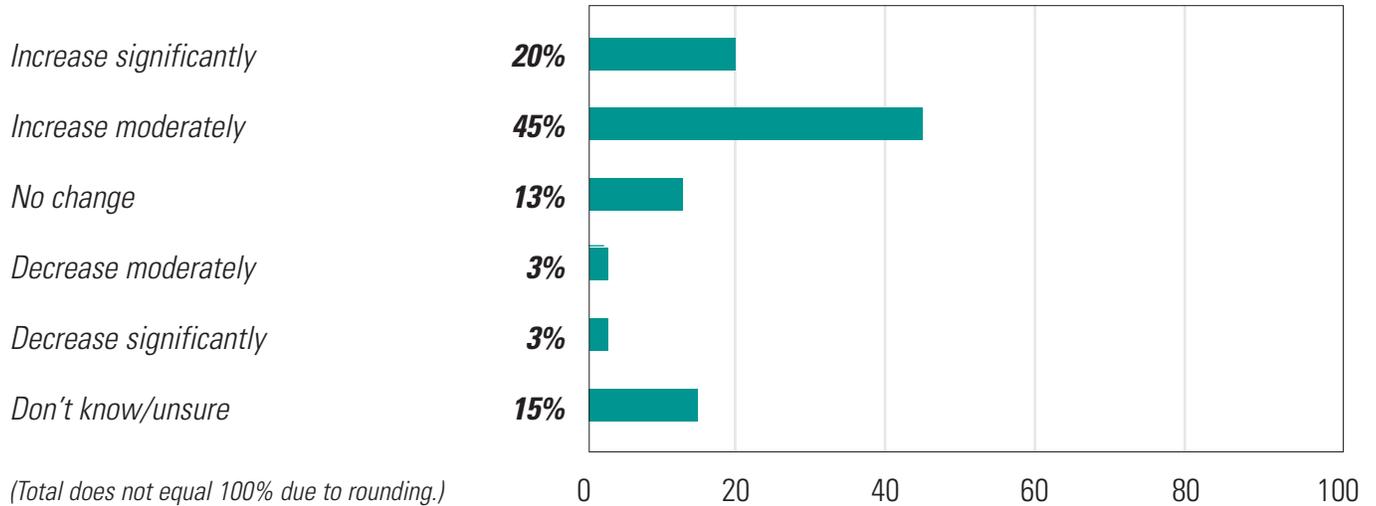
## Figure 3: Percentage of Unstructured Data Within Enterprises



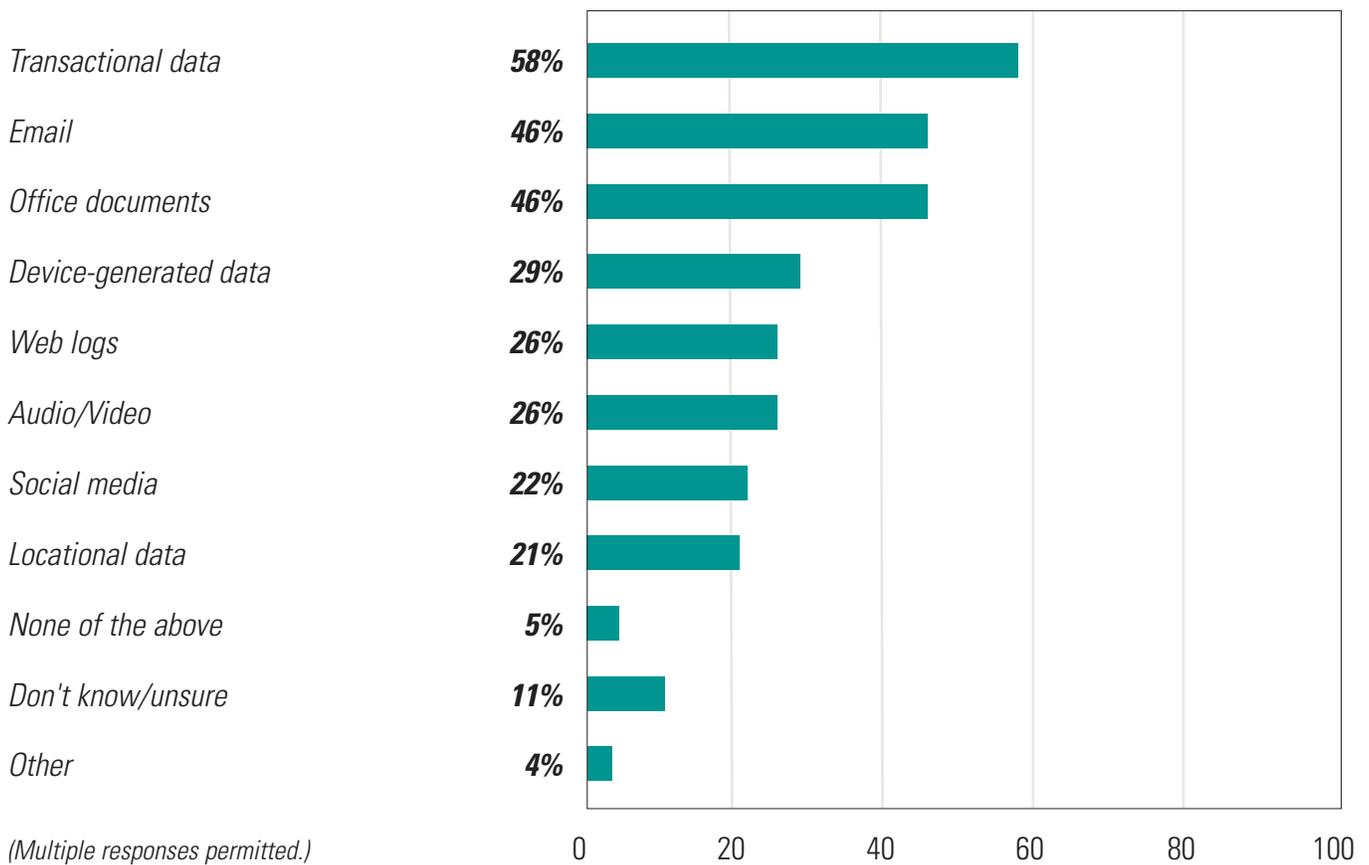
**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 4: How Amount of Unstructured Data Within Enterprises Will Change Within 3 Years



## Figure 5: Data Types Driving Need for Big Data Technologies



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 6: IT Infrastructure and Database Systems Adequate for Managing Data?

|                          | <i>Now</i> | <i>In 3 Years</i> |
|--------------------------|------------|-------------------|
| <b>Yes</b>               | <b>28%</b> | <b>19%</b>        |
| <b>Mostly</b>            | <b>37%</b> | <b>36%</b>        |
| <i>Somewhat</i>          | <b>21%</b> | <b>21%</b>        |
| <i>No</i>                | <b>7%</b>  | <b>12%</b>        |
| <i>Don't know/unsure</i> | <b>6%</b>  | <b>11%</b>        |

(Totals do not equal 100% due to rounding.)

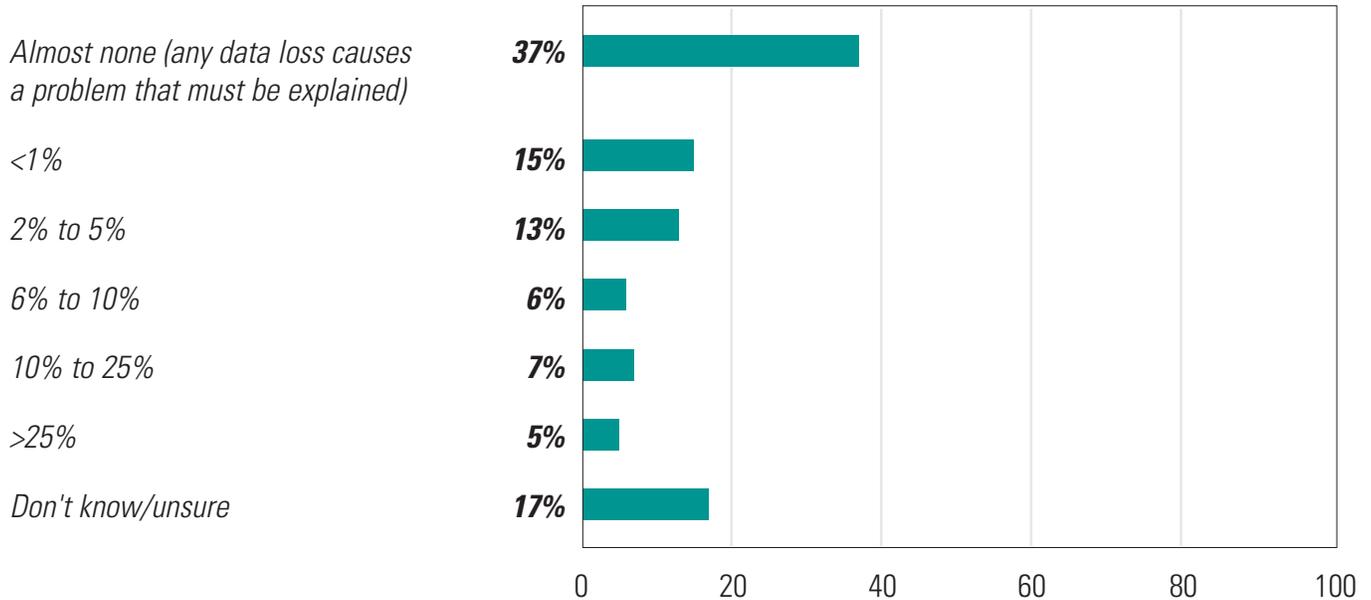
## Figure 7: IT Infrastructure and Database Systems Adequate for Managing Data—By Data Type

(Percentage answering “yes”)

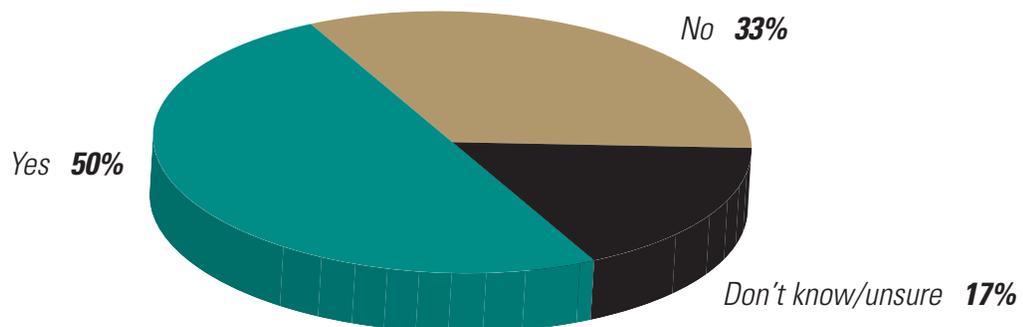
|                        | <i>Now</i> | <i>In 3 Years</i> |
|------------------------|------------|-------------------|
| <i>Big Data</i>        | <b>21%</b> | <b>11%</b>        |
| <i>Small Footprint</i> | <b>32%</b> | <b>21%</b>        |

(*Big Data* organizations defined as those reporting >500TB and >25% unstructured data stores. “*Small Footprint*” defined as those reporting <50TB and <25% unstructured data.)

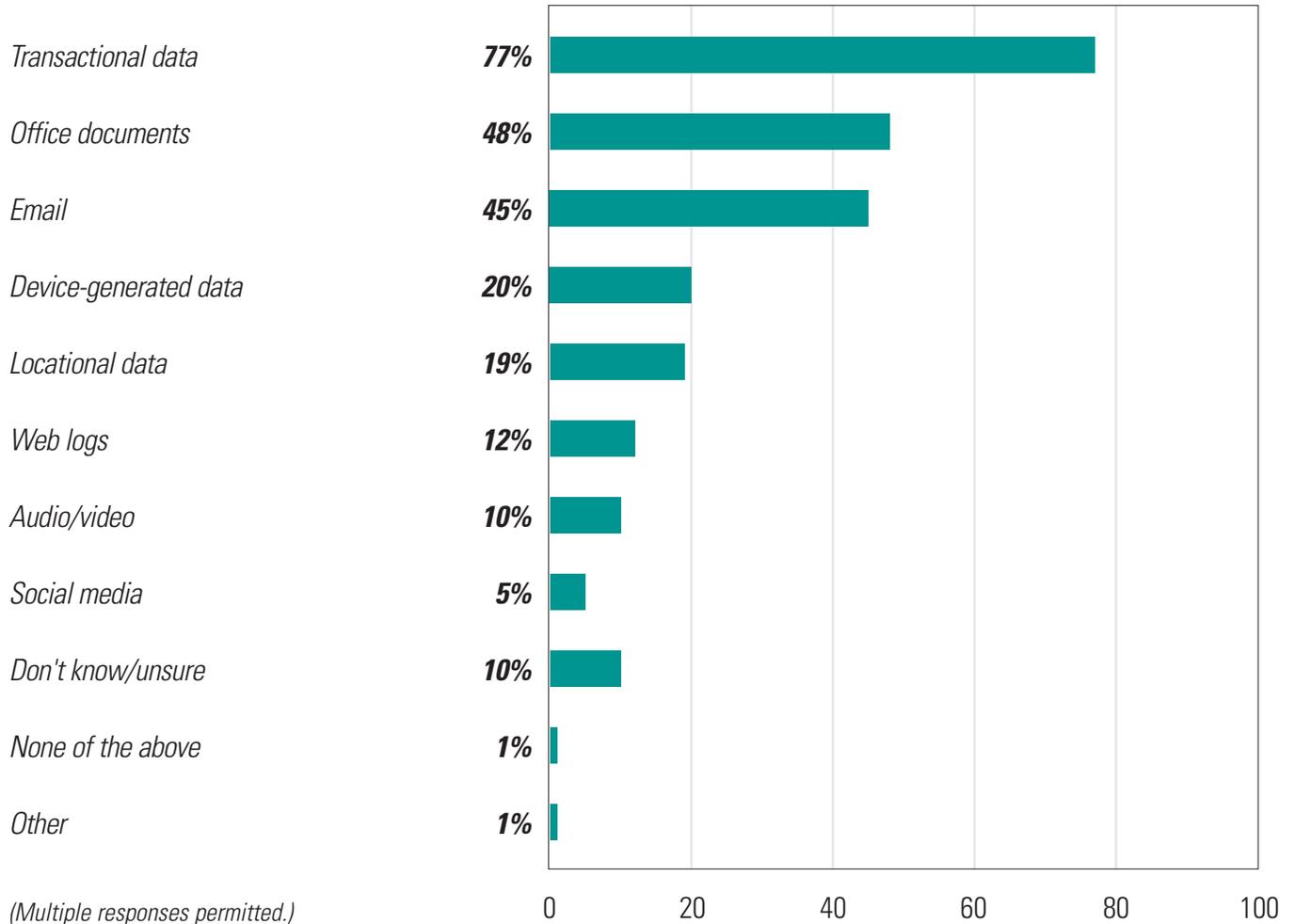
## Figure 8: How Much Data Can Be Lost Without Serious Harm?



## Figure 9: Does Tolerance for Data Loss Vary by Data Type?



## Figure 10: Data to be Guarded Against Loss



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## BUSINESS DRIVERS

While many organizations are still struggling to understand the business value of Big Data, more than 50% consider it to be extremely or very important to their business. The greatest opportunities Big Data offers are in competing more effectively (40%) and growing business revenue streams (32%). Most Big Data initiatives currently come out of the IT department, but the first business applications are being seen in marketing and sales—an area typically already comfortable with data analytics.

In many cases, Big Data is not something businesses have learned to capitalize on as of yet. While 55% say Big Data is “extremely” or “very” important to their businesses, there is still a large remaining segment of respondents who are not sure of its importance. (See Figure 11.) Ultimately, the goal is to develop an analytics strategy that takes advantage of the Big Data resources that are now being captured and maintained within enterprises.

The benefits of being able to leverage Big Data into business value are far-reaching, and respondents are generally optimistic as to what it can provide for them. The initial business case starts with competitive differentiation and marketing. The largest number of respondents see Big Data as an opportunity to compete more effectively within their markets (41%), as well as better understand their customers (40%). At least one-third of respondents also look to Big Data as providing better opportunities to grow business revenue streams, as well as lower business operational costs. (Figure 12.)

Getting the attention and support of the C-suite for such efforts often can be challenging, respondents report. What are the main barriers to using Big Data within respondents’ companies? The leading issue is lack of budget, cited by 40%. About a third of respondents also say that Big Data simply is not yet a business management priority. Technology issues also get in the way, as indicated by close to a third of respondents. Often, organizations attempting to manage Big Data have diverse environments in which it is difficult to take advantage of Big Data. Data may be highly distributed, requiring a Big Data strategy that is built on a combination of different storage devices and different databases that store structured data. (See Figure 13.)

While close to half of the respondents, 46%, are not yet sure how Big Data will be incorporated into their BI analysis, about one third of the respondents, 32%, pre-process Big Data, then load it into their data warehouse for integrated analysis. This suggests that the majority of respondents who are using Big Data find the greatest value in integrating the emerging unstructured data world with existing relational data environments. Another 14% say they conduct Big Data analysis separately from traditional enterprise analysis. (See Figure 14.)

IT and business domain experts within organizations need to work together to develop and govern Big Data opportunities,

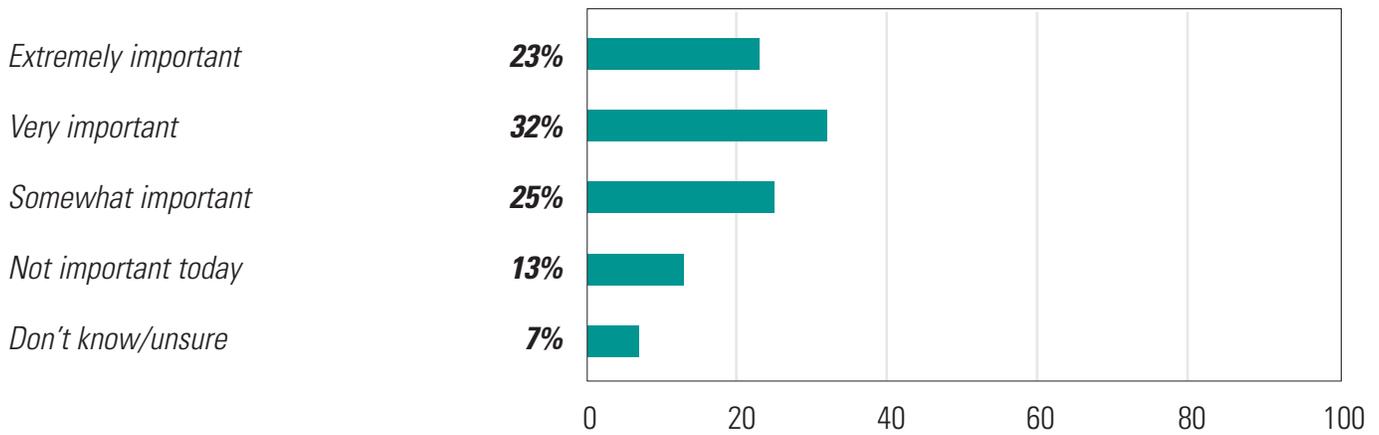
as noted by one respondent, an IT executive with a consulting firm. “We already have as close a relationship with management as is possible,” says the respondent. “We intend to keep it that way by doing a great job on Big Data, but no matter what we do in this area where terms and ideas are still the subject of daily definition, someone else could do it better. There is no standard of achievement that will be safely considered good. We have no idea what the percentage of the data flying past [is] us good enough to capture. Understanding the potential benefits and liabilities of capturing a wide range of data beyond traditional transactions is an open-ended subject that will go on for a century.”

In a majority of cases, 51%, Big Data projects originate in the IT department. (See Figure 15.) Among respondents who report Big Data projects originating within their organizations’ line-of-business departments, a majority, 54%, report that marketing and sales tend to generate Big Data projects. This is the part of the business that historically has developed analytics against data warehouse information, from customer segmentation to market-basket analysis. At least a third of respondents also indicate that some core areas of backbone business operations, including administration and production, also are engaging with Big Data. (See Figure 16.)

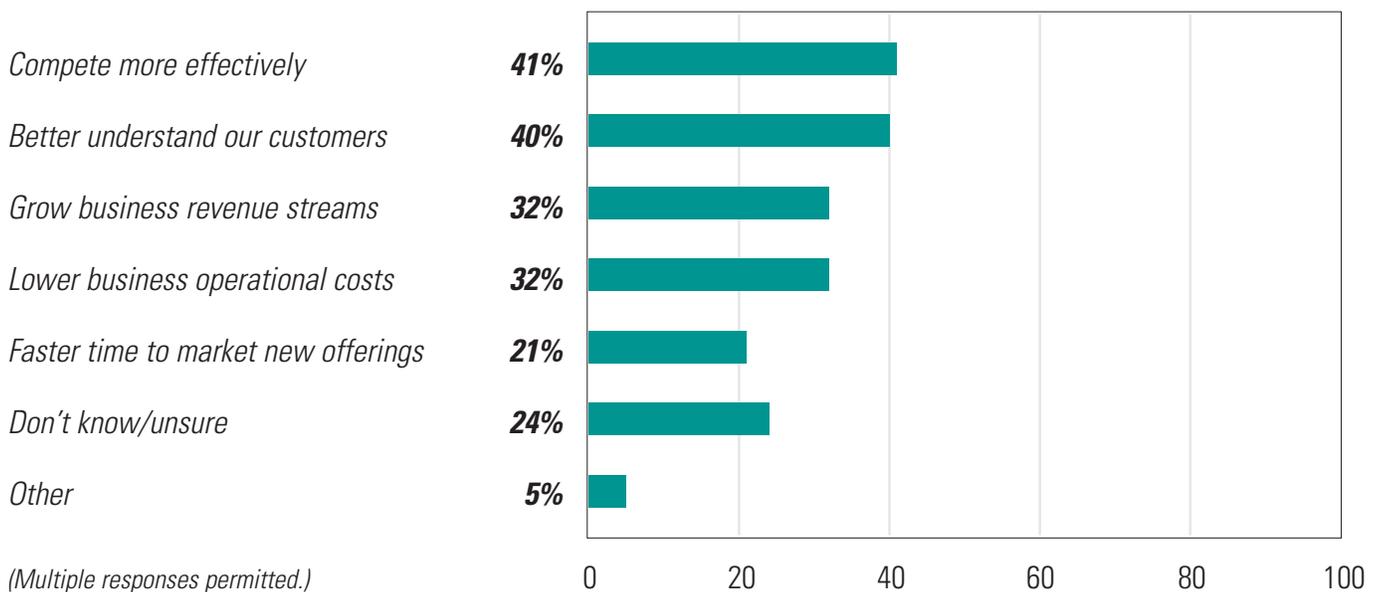
Who in respondents’ organizations influences purchasing decisions for Big Data projects? Half of respondents, 49%, say their IT departments have the greatest influence. Likewise, 48% say their CIOs have purchasing authority. Ultimately, however, the final dollar decision is up the C-level business executives. At this stage, only 30% say their IT departments can sign off on final expenditures. (See Figure 17.)

Then, when it comes time for implementation, responsibility is handed back to members of the IT department—who, along with enterprise data architects and BI/analytics teams in some cases, are responsible for making it all work. (See Figure 18.) In larger organizations, these responsibilities tend to be assumed by specialized BI/analytics teams and enterprise data architects, while small organizations put their IT managers in charge. Interestingly, 18% of both small and large organizations report they have data scientists helping to lead the way. (See Figure 19.)

## Figure 11: Importance of Big Data to Business



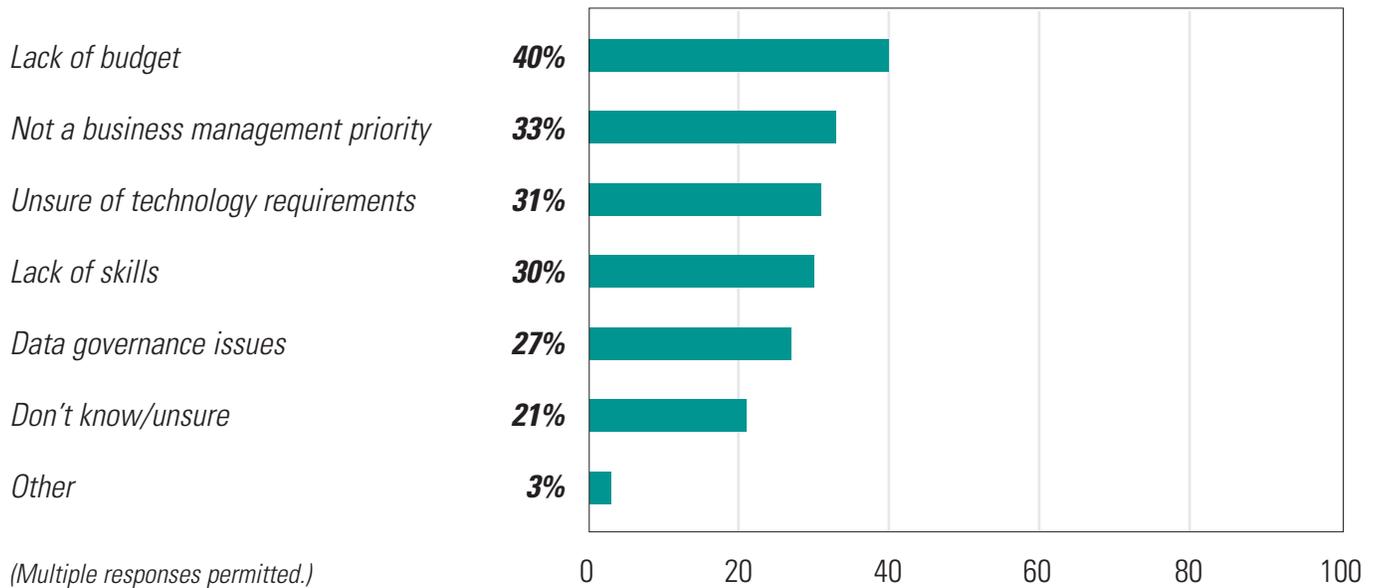
## Figure 12: Leading Big Data Business Opportunities



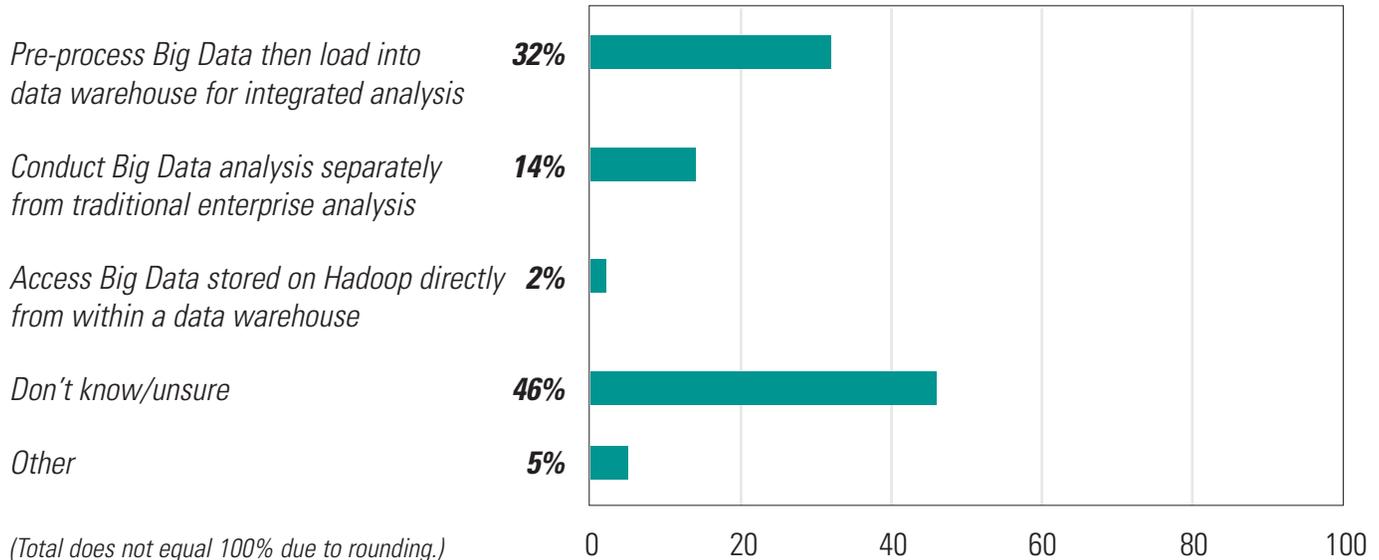
**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

### Figure 13: Main Barriers to Using Big Data Within Companies



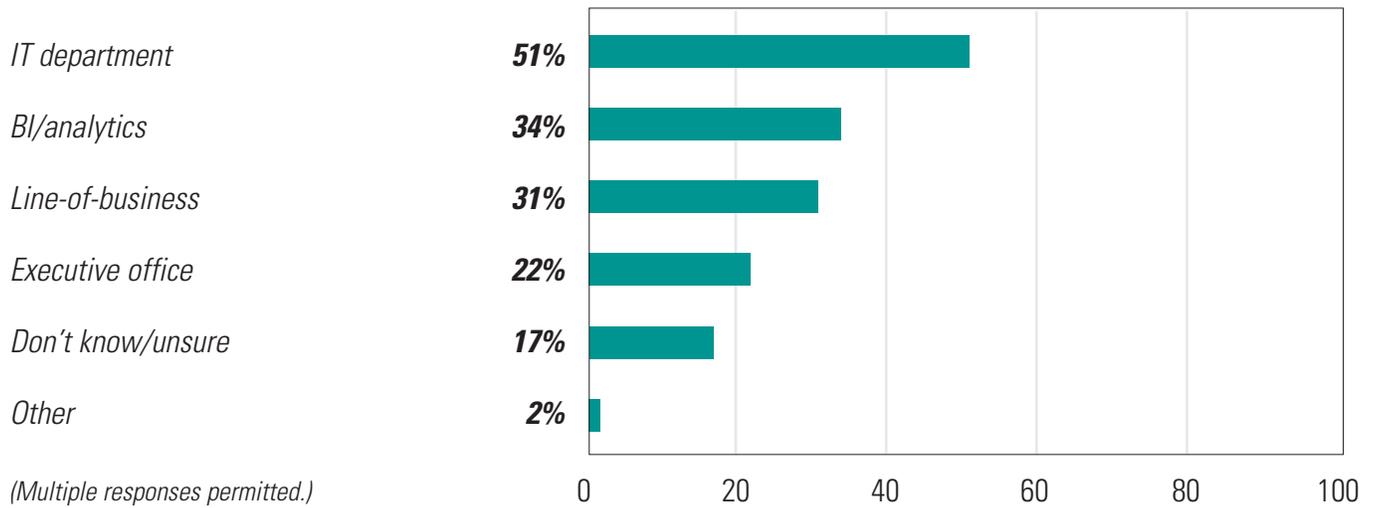
### Figure 14: How Big Data is Integrated With BI Applications



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 15: Departments from Which Big Data Projects Originate

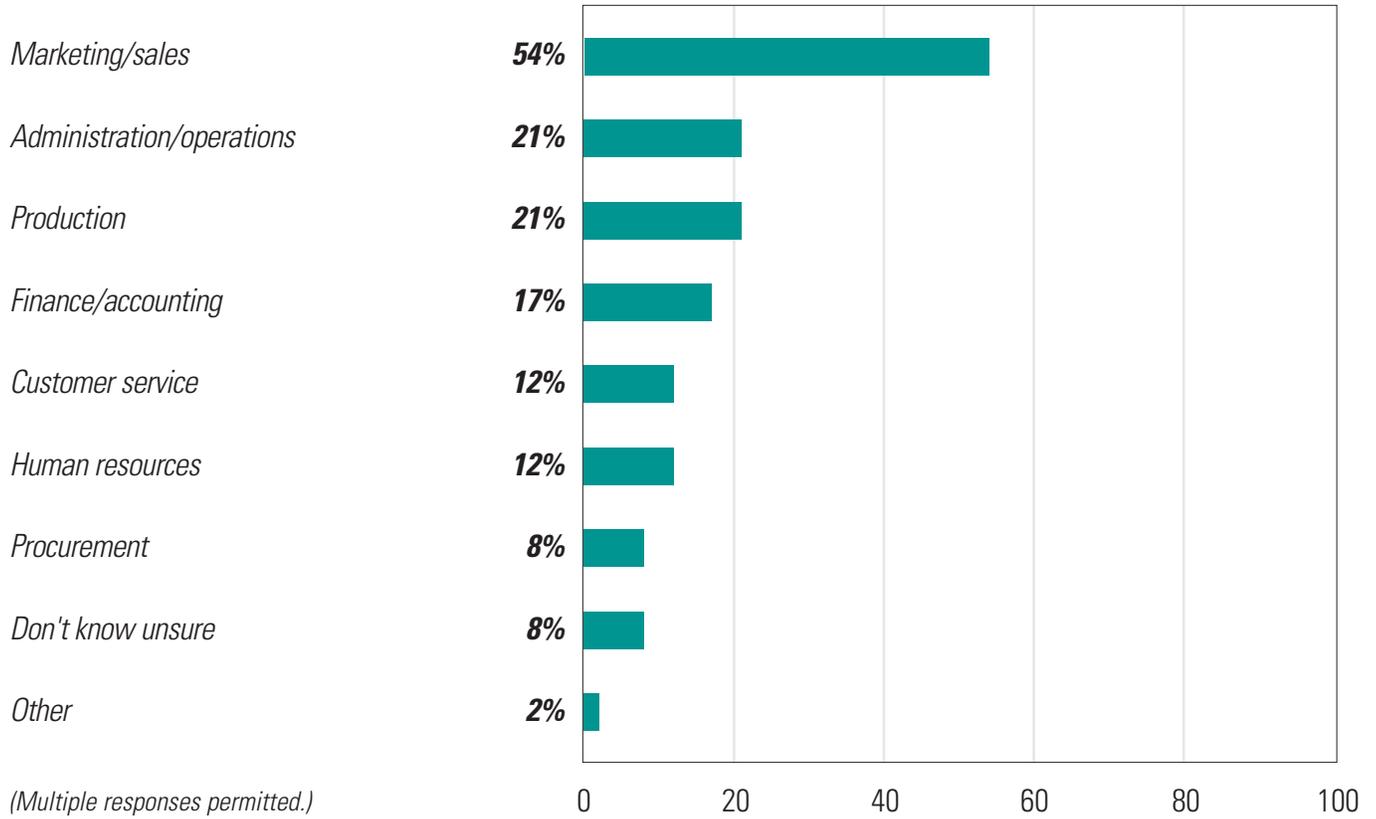


**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 16: Business-Side Driver of Big Data Initiatives

(Among respondents indicating the Big Data projects originate within their line-of-business departments)



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

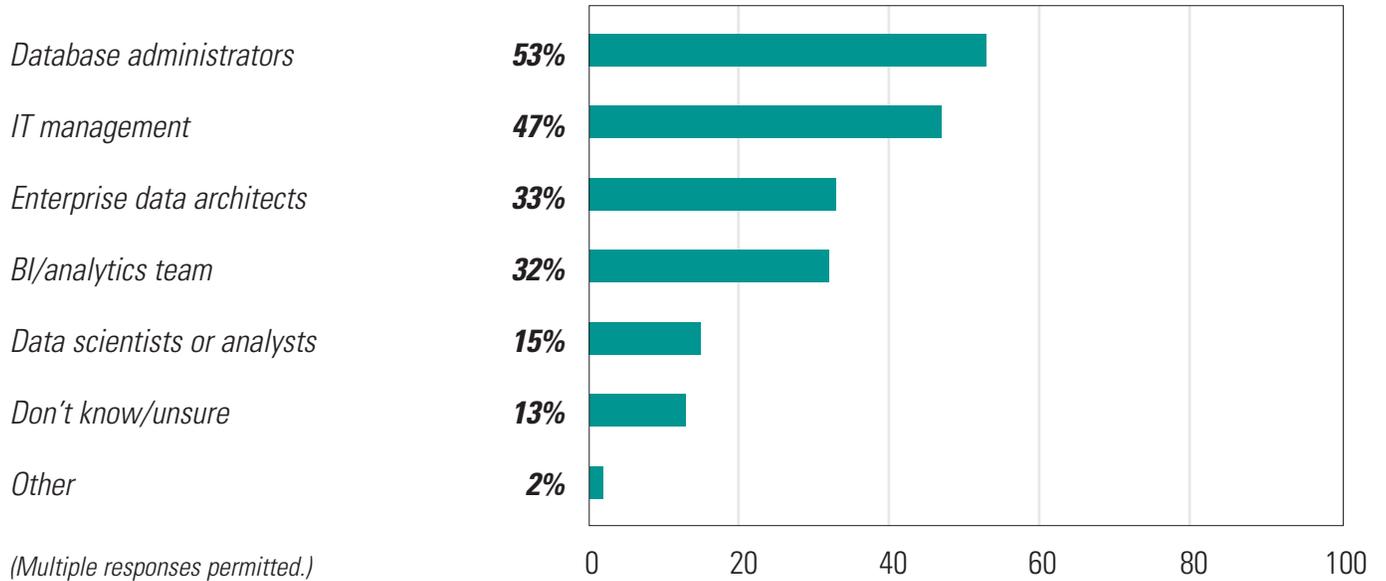
Data collection and analysis performed with SurveyMethods.

## Figure 17: Who in Organization Influences Purchasing Decisions for Big Data Projects

|   | <i>Influences</i> | <i>Final Decision</i> |
|---|-------------------|-----------------------|
| <i>IT management</i>                            | <b>49%</b>        | <b>30%</b>            |
| <i>CIO</i>                                      | <b>48%</b>        | <b>48%</b>            |
| <i>Executive business mgmt. (CEO, CFO, COO)</i> | <b>37%</b>        | <b>6%</b>             |
| <i>Enterprise data architects</i>               | <b>34%</b>        | <b>6%</b>             |
| <i>Data scientists or analysts</i>              | <b>16%</b>        | <b>3%</b>             |
| <i>Database administrators</i>                  | <b>30%</b>        | <b>3%</b>             |
| <i>BI/analytics team</i>                        | <b>25%</b>        | <b>12%</b>            |
| <i>Line-of-business management</i>              | <b>22%</b>        | <b>12%</b>            |
| <i>Don't know/unsure</i>                        | <b>14%</b>        | <b>13%</b>            |
| <i>Other</i>                                    | <b>0%</b>         | <b>1%</b>             |

*(Multiple responses permitted.)*

## Figure 18: Who in Organization is Responsible for Implementing and Managing Big Data Solutions



## Figure 19: Who in Organization is Responsible for Implementing and Managing Big Data Solutions—By Company Size

|                             | <1,000 employees | 1,000–10,000 employees | >10,000 employees |
|-----------------------------|------------------|------------------------|-------------------|
| IT management               | 54%              | 47%                    | 44%               |
| Enterprise data architects  | 31%              | 30%                    | 38%               |
| BI/analytics team           | 17%              | 30%                    | 54%               |
| Data scientists or analysts | 18%              | 11%                    | 18%               |
| Database administrators     | 50%              | 51%                    | 62%               |

(Multiple responses permitted.)

**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## BIG DATA MANAGEMENT TOOLS AND PLATFORMS

**Hadoop—often seen as the cornerstone of many Big Data efforts—has yet to catch on within most Oracle environments. This suggests that many organizations are just starting to address Big Data storage requirements. Adoption of Hadoop, in fact, is expected to double over the coming year.**

As this survey shows, respondents have growing data stores, and an incredible variety of unstructured files now coming under their purview. Business end users—especially marketing and sales departments—are anxious to get at this data to better understand where to steer the business. The goal is that the rise of Big Data can also create a lack of cohesion within enterprises, many of which rely on siloed architectures. The challenge is to be able to capture the data coming from many sources at high speed, process that data, put it into highly available storage, and be able to use that data in data analytics or visualization applications. But data tends to end up in different stores or different silos, making it difficult for decision makers to get at and discover the hidden relationships that may exist within that data.

The most prevalent data technologies in use at respondents' companies include relational databases such as Oracle (92%) and business intelligence tools (47%). Interestingly, the two technologies most closely associated with Big Data—Hadoop/MapReduce and NoSQL databases—are only seen among a minority of respondents at this time (7% and 11%, respectively). However, looking ahead, there will be increased adoption, with usage of Hadoop/MapReduce more than doubling in the next year from 7% to 16% and NoSQL database usage growing from 11% to 15%. Relational databases are still the top technology choice going forward. (See Figure 20.)

Hadoop is an option for bringing a large quantity of unstructured data into a manageable file system that can be readily ingested by applications across the organization. Hadoop, an Apache project, enables applications to run across large arrays of nodes, accessing petabytes' worth of data. The Hadoop framework is built on the Hadoop Distributed File System (HDFS), which stores files across storage nodes in a Hadoop cluster. About 13% of respondents have deployed or are in the process of implementing or piloting Hadoop at this time, with an additional 22% considering adoption of the open source framework at some point in the future—and, as noted above, 16% plan to adopt the framework over the coming year. (See Figure 21.)

Hadoop users in the survey tend to have greater data challenges than those not considering the framework. More than four-fifths of respondents with Hadoop in place, or being implemented or piloted, for example, see Big Data as an integral to their business, versus only 48% of non-Hadoop sites. (See Figure 22.)

In addition, Hadoop is four times as likely to be implemented at organizations with large stores of unstructured data (comprising the majority of their data store). (See Figure 23.) In addition, only 15% of those using or planning to use Hadoop are satisfied with their corporate data infrastructure (versus 31% of non-Hadoop sites), making a clear-cut case for their implementation of the open-source framework to better manage Big Data. (See Figure 24.)

For respondents who are already working with Hadoop, the most common function is for managing ad hoc queries (17%), followed by data mining (14%), staging data (13%) and as a transformation engine (11%). The percentage employing Hadoop for ad hoc queries will rise to 31% within the next 24 months, and data mining will rise to 35%, making these the most prevalent use cases for Hadoop adoption in the near future. (See Figure 25.)

The largest segment of current Hadoop users, 11%, say their Hadoop clusters are still relatively small, not exceeding 100TB. (See Figure 26.)

For those respondents using Hadoop, most anticipate rapid growth. About half of respondents in this subgroup, in fact, expect the size of their clusters to increase significantly. (See Figure 27.)

Tuning cluster performance, along with application development, are the two greatest challenges faced by respondents working to deploy Hadoop, the survey finds. (See Figure 28.)

About 20% of organizations in the survey use open source technologies within their data environments, and 23% expect to continue expanding their use of these solutions over the coming year. In addition to established commercial offerings from big IT vendors, open source tools now have the usability features that make them viable for the enterprise. The ecosystem that's developed around these open source applications—cloud-based offerings, third-party integration, and developer support—has further established them as a competitive factor. Today's buyer has a wealth of purchasing options, from roll-your-own applications built on open source tools to cloud-based offerings to solutions sourced from IT vendors.

How do respondents support open source tools within their organizations? For the most part, respondents pursue measures most commonly associated with the open source world: pulling resources from the web. The largest segment, 36%, rely on ad hoc



internal resources and community support, while 35% look to dedicated community resources. (See Figure 29.)

With today's new workloads come large amounts of data, such as log records as well as transaction-related information. Big Data may offer significant opportunities for analysis and insights on a scale never seen before in business, but, as reported earlier, most respondents do not feel their existing data infrastructure is ready for the job. The success of Big Data analytical efforts depends on how well organizations can store, manage and render actionable information from data that is streaming in from users and systems from both within and outside.

The question is, will companies turn to cloud technologies to address these challenges? Public cloud is not yet playing a major role in Big Data initiatives. Only 8% of respondents currently use a cloud provider to support Big Data initiatives. Another 21% are considering taking such steps, however. (See Figure 30.)

However, those respondents using or implementing Hadoop are far more likely to be embracing both public and private cloud networks. The connection between cloud adoption and Big Data isn't clear-cut yet. But the flexibility of being able to store and manage data in the cloud that is ingested by Hadoop—

and perhaps access the framework itself through a cloud site—may be an important factor in the difference in adoption rates. (See Figure 31.)

Overall, for all needs, private cloud is the most common configuration, indicated by 38% of respondents. About three out of 10 employ hosted solutions from their software vendors, and about 12% overall use public cloud services in some capacity, which may or may not include Big Data management. (See Figure 32.)

Private clouds are most commonly seen within larger organizations—in this survey, 48% of those respondents in organizations with more than 10,000 employees now support private cloud implementations. Larger organizations are also twice as likely to be adopting public cloud services than their small-business counterparts. (See Figure 33.)

There are a number of advantages respondents see to using a cloud provider to support Big Data initiatives. One-third say such providers offer more cost-effective options. Faster time to value is another option, as indicated by 27%. (See Figure 34.) However, there are challenges to Big Data cloud as well. A majority, 60%, cite security concerns, while 40% are concerned with quality control. (See Figure 35.)

## Figure 20: Data Management Technologies in Use or Planned for Adoption

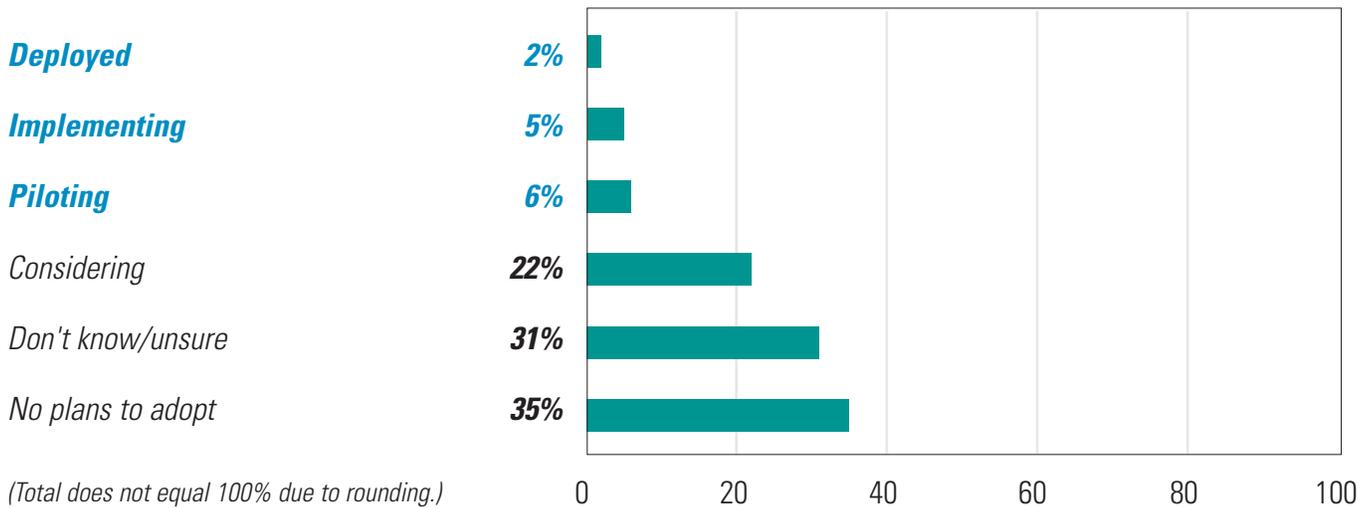
|   | <i>Currently</i> | <i>Within 12 Months</i> |
|---|------------------|-------------------------|
| <i>Relational databases (e.g., Oracle, DB2, etc.)</i> | <b>92%</b>       | <b>58%</b>              |
| <i>Business intelligence tools</i>                    | <b>47%</b>       | <b>39%</b>              |
| <i>Other open source technologies</i>                 | <b>20%</b>       | <b>23%</b>              |
| <i>NoSQL databases</i>                                | <b>11%</b>       | <b>15%</b>              |
| <i>R</i>  | <b>3%</b>        | <b>5%</b>               |
| <i>Don't know/unsure</i>                              | <b>5%</b>        | <b>20%</b>              |
| <i>Other</i>  | <b>3%</b>        | <b>3%</b>               |

*(Multiple responses permitted.)*

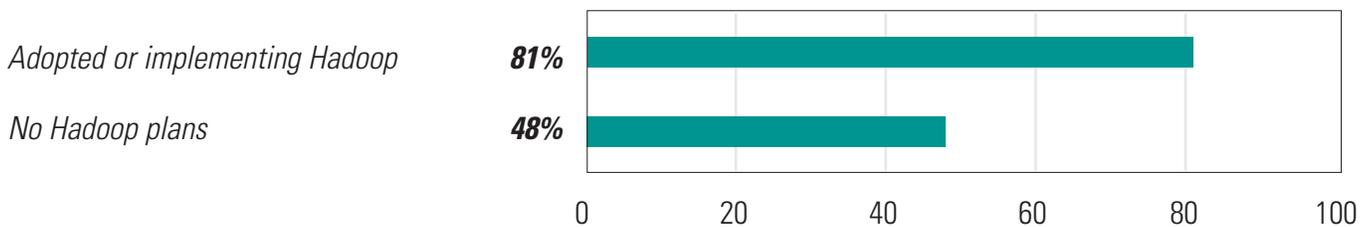
**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

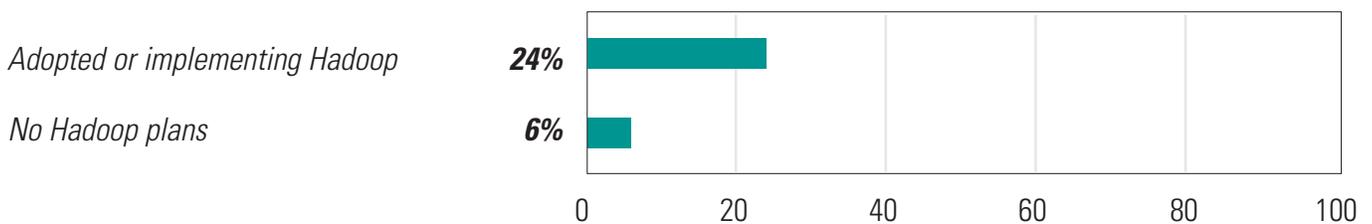
## Figure 21: Hadoop Adoption Cycle



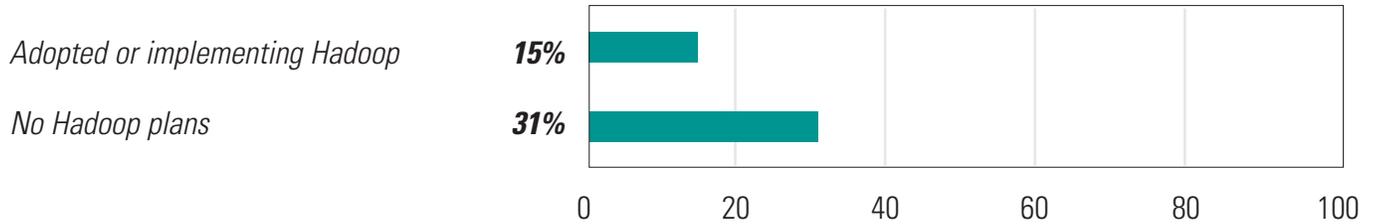
## Figure 22: Hadoop Users Are More Likely to See Big Data as “Extremely” or “Very Important” to Their Business



## Figure 23: Hadoop Users Are More Likely to Be At Mainly Unstructured Data Sites



## Figure 24: Hadoop Users Are More Less Likely to Consider Their Legacy Data Infrastructure to be “Adequate”

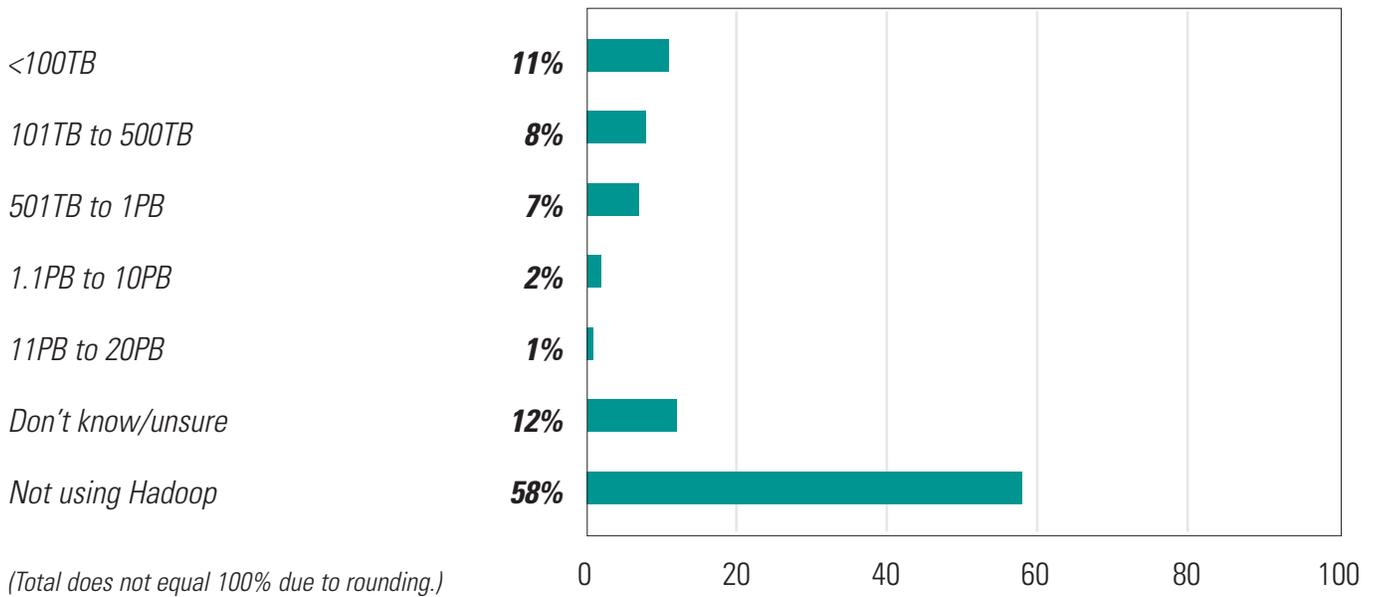


## Figure 25: How Hadoop Is and Will be Used

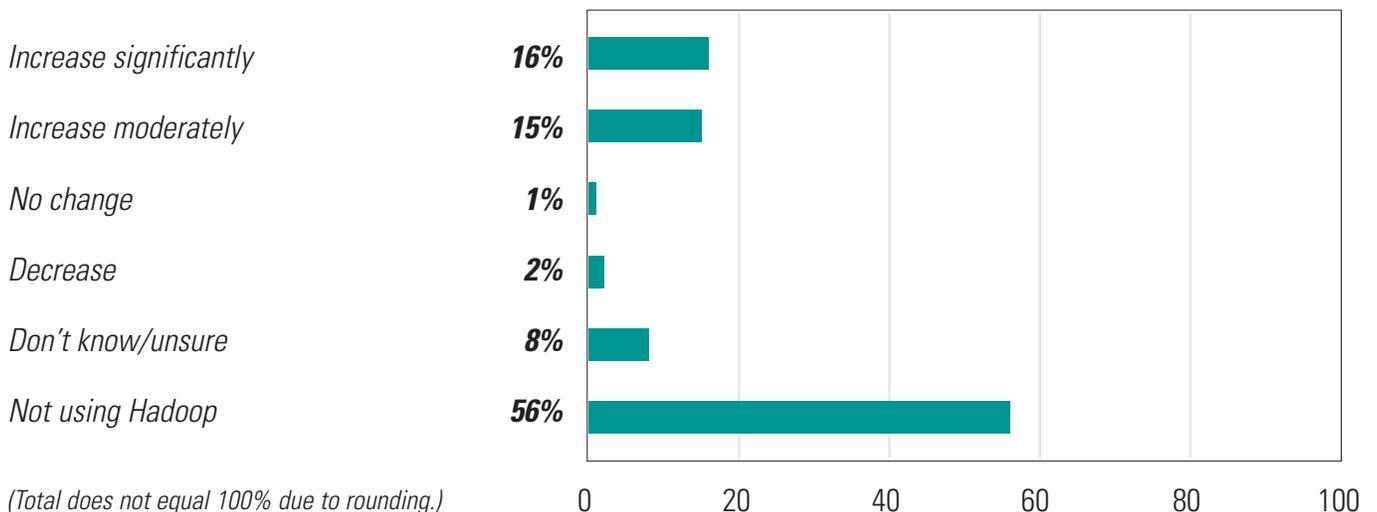
|                       | Currently | Within 12 Months |
|-----------------------|-----------|------------------|
| Ad hoc queries        | 17%       | 31%              |
| Data mining           | 14%       | 35%              |
| Staging area          | 13%       | 24%              |
| Transformation engine | 11%       | 27%              |
| Scheduled reports     | 10%       | 20%              |
| Online data archive   | 7%        | 20%              |
| Don't know/unsure     | 13%       | 14%              |
| Not using Hadoop      | 58%       | 28%              |
| Other                 | 0%        | 2%               |

(Multiple responses permitted.)

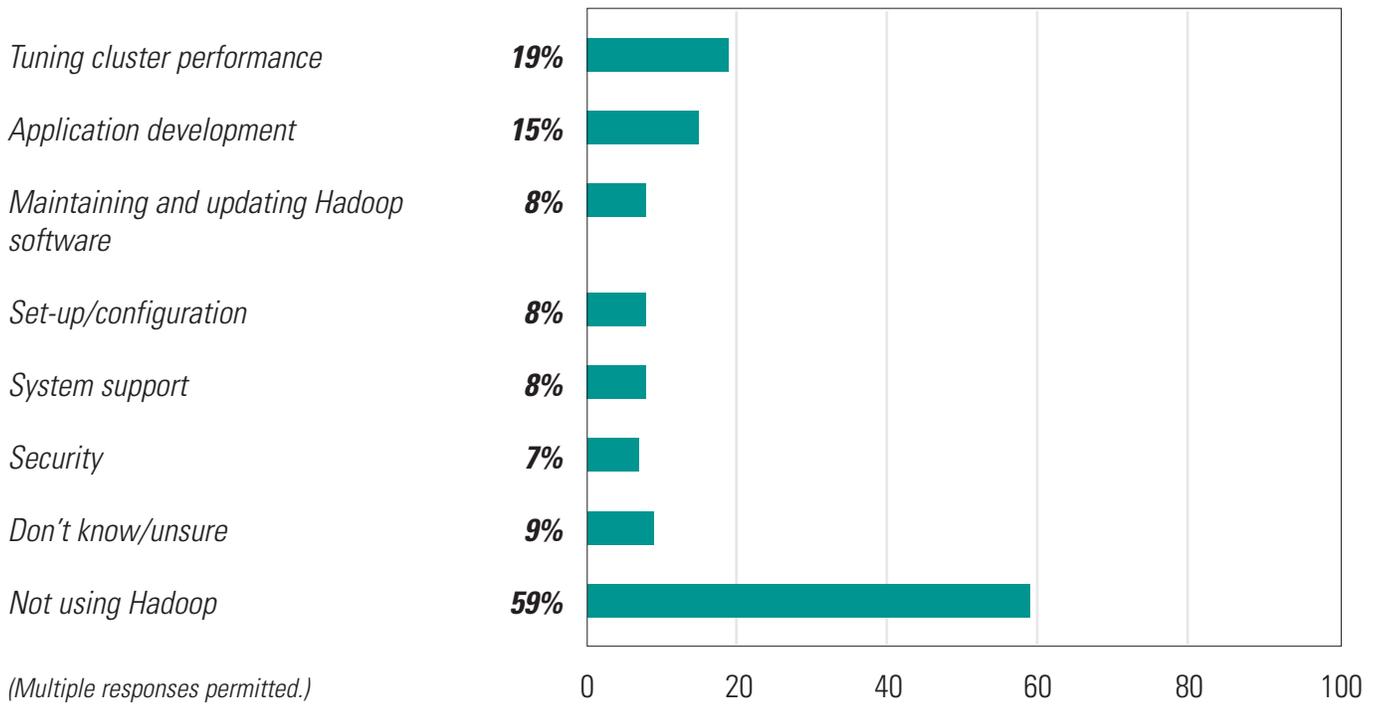
## Figure 26: Size of Hadoop Clusters



## Figure 27: Expected Growth of Hadoop Cluster Over Next 3 Years



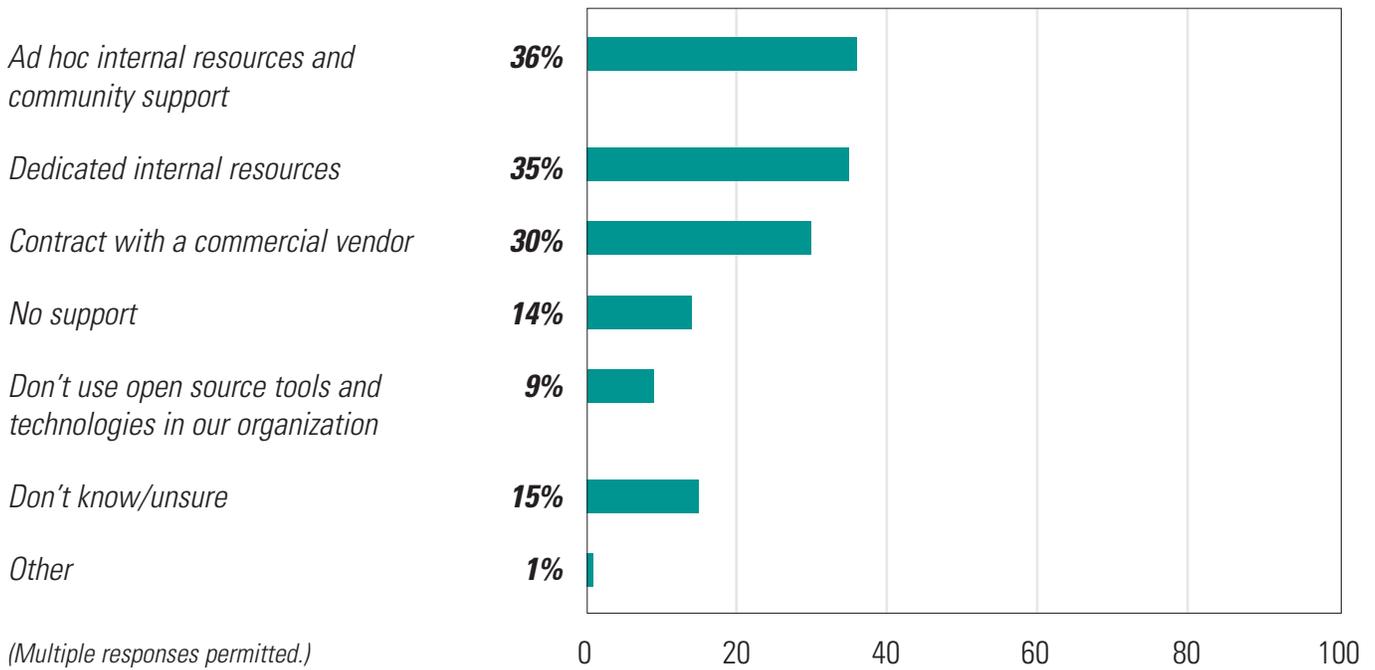
## Figure 28: Challenges to Managing Hadoop Clusters



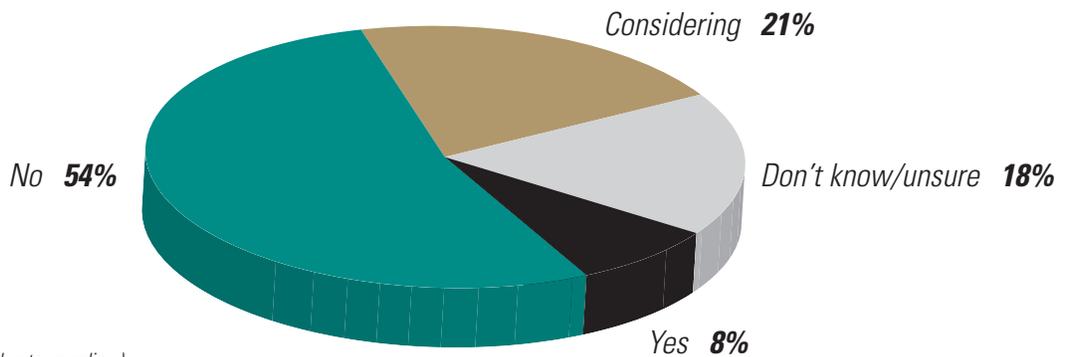
**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

**Figure 29: How Open Source Tools are Supported**



**Figure 30: Using a Cloud Provider to Support Big Data?**



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

### Figure 31: Hadoop Users More Likely to Embrace the Cloud

|                                       | <i>Public Cloud</i> | <i>Private Cloud</i> |
|---------------------------------------|---------------------|----------------------|
| <i>Adopted or implementing Hadoop</i> | <b>39%</b>          | <b>55%</b>           |
| <i>No Hadoop plans</i>                | <b>7%</b>           | <b>29%</b>           |

### Figure 32: Types of Cloud Services Currently Used for All Purposes, Including Big Data

|   | <i>Currently</i> | <i>Next 12 Months</i> |
|---|------------------|-----------------------|
| <i>Private cloud services/applications</i>              | <b>38%</b>       | <b>37%</b>            |
| <i>Hosted solutions</i>                                 | <b>29%</b>       | <b>30%</b>            |
| <i>Public cloud</i>                                     | <b>12%</b>       | <b>8%</b>             |
| <i>Cloud-based database system</i>                      | <b>8%</b>        | <b>18%</b>            |
| <i>We do not use or plan to use cloud-based service</i> | <b>30%</b>       | <b>15%</b>            |
| <i>Don't know/unsure</i>                                | <b>14%</b>       | <b>28%</b>            |

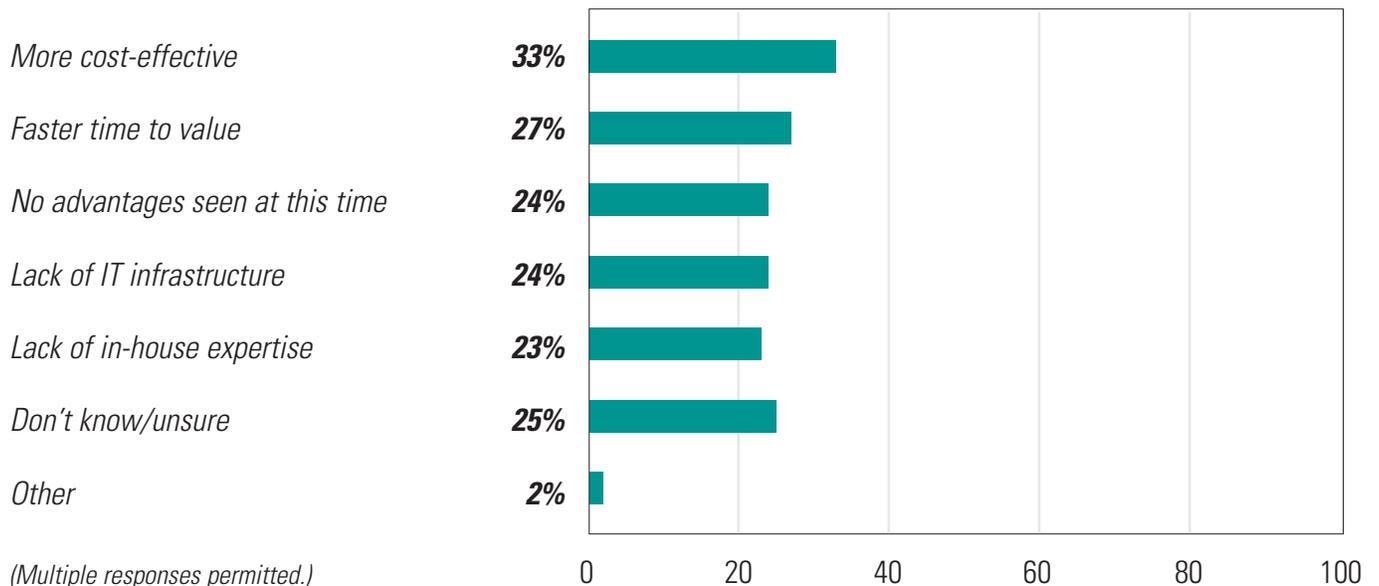
*(Multiple responses permitted.)*

### Figure 33: Types of Cloud Services Currently Used for All Purposes, Including Big Data—By Company Size

|  | <1,000 employees | 1,000–10,000 employees | >10,000 employees |
|--|------------------|------------------------|-------------------|
| Private cloud services/applications                      | 32%              | 37%                    | 48%               |
| Hosted solutions   | 30%              | 25%                    | 38%               |
| Public cloud   | 8%               | 12%                    | 16%               |
| Cloud-based database system                              | 4%               | 8%                     | 13%               |
| We are a public cloud provider or hosting service        | 4%               | 3%                     | 10%               |
| We do not use or plan to use/provide cloud-based service | 4%               | 3%                     | 10%               |
| Don't know/unsure  | 35%              | 29%                    | 26%               |

(Multiple responses permitted.)

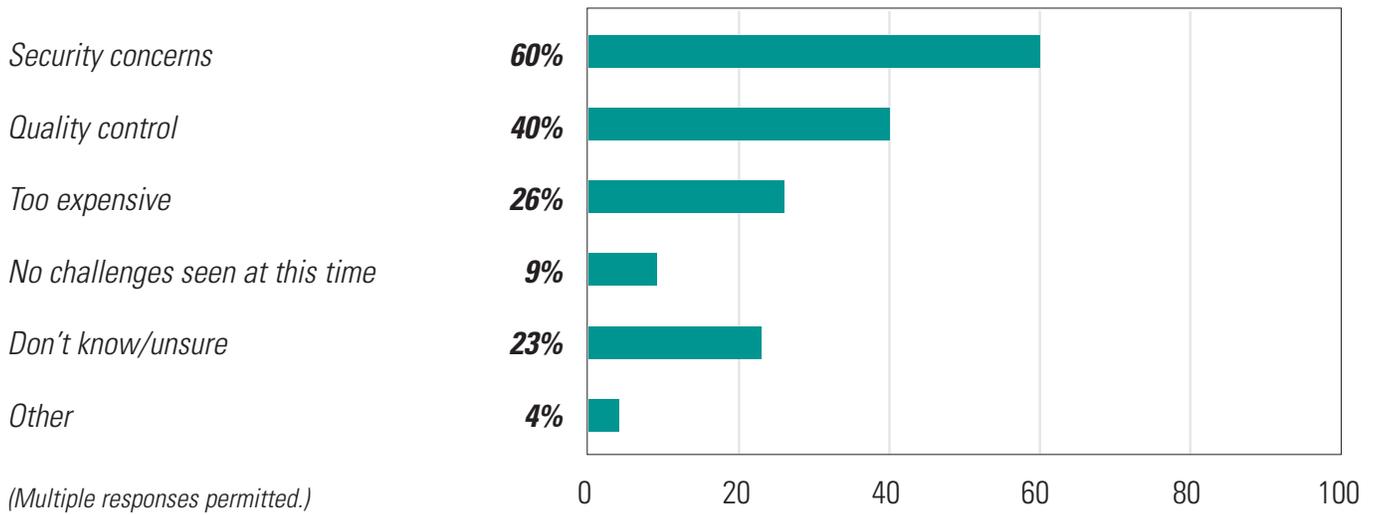
### Figure 34: Advantages of Big Data Cloud



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 35: Challenges With Big Data Cloud



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

---

## IOUG RECOMMENDS

**Tapping the business potential of Big Data requires the ability to access and integrate information at any scale, from any source. This calls for effective management and deployment of data platforms and processes that can deliver results in areas such as customer relationship management, sales performance, operational efficiency, and business agility. IOUG recommends the following approaches to meet these burgeoning needs for availability in fast-growing and increasingly complex Big Data environments:**

**Develop a business case.** Big Data may be a headache at times for data professionals, but for the business, it represents Big Opportunities. Never before have decision makers had access to insights from so many parts of the business at once. However, data managers and professionals need to help business decision makers “filter” through the massive amounts of data, and noise, to identify key nuggets of pertinent information. This requires working with end users to identify what types of data will have the greatest impact.

**Get business buy-in and support.** Big Data analytics only deliver value if they have the support and input of the business. The business needs to determine what data needs to be made available to analytic platforms, and what data isn’t as essential. In the survey, many respondents see this as a way to strengthen the relationship between IT and the business.

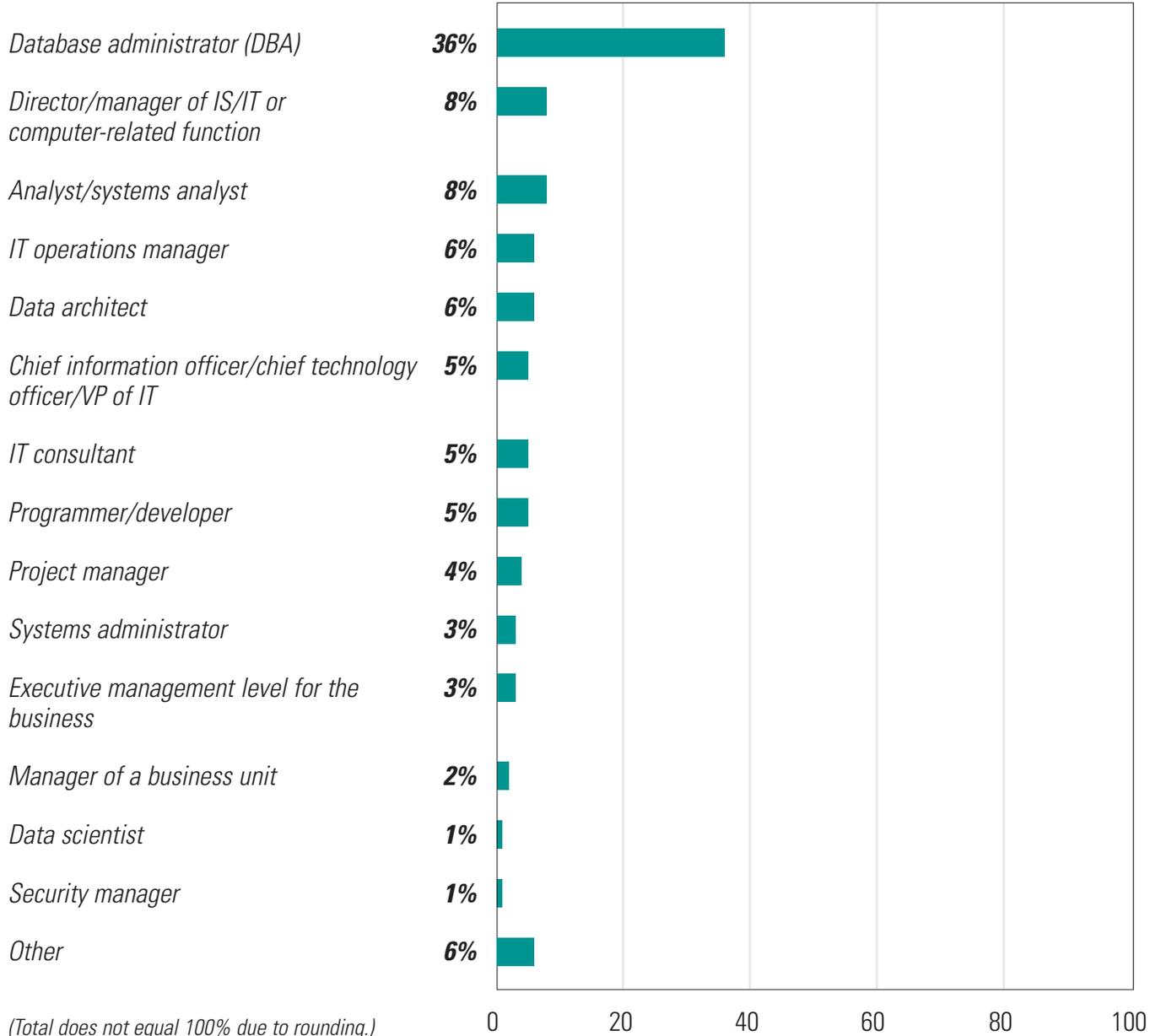
**Develop an integration strategy between unstructured and traditional enterprise data.** The levels of unstructured data flowing into enterprises are growing rapidly, and businesses are seeking ways to develop insights from this data. Today’s generation of solutions provide platforms to bring both unstructured and structured transactional data together into a common environment.

**Develop an integrated information management lifecycle strategy.** Some data needs to reside online and be quickly accessible to end users, while other forms of data can be less accessible and stored in archive or backup systems. More Big Data flowing into organizations may place greater stress on current infrastructure, affecting application performance and availability. Data should be moved to tiered storage systems as part of an organization’s entire information management lifecycle processes.

The findings from this latest IOUG member survey show that many organizations already have massive data stores in terms of volume and variety. The challenge of maintaining this level of online data access is increasing as Oracle and other database sites are being inundated with the management of large volumes of enterprise data, both structured and unstructured. This increase in volume, velocity, and variety of data not only presents management and storage challenges, but also offers great opportunities for businesses to expand their capability to innovate, better serve customers, and make more insightful decisions.

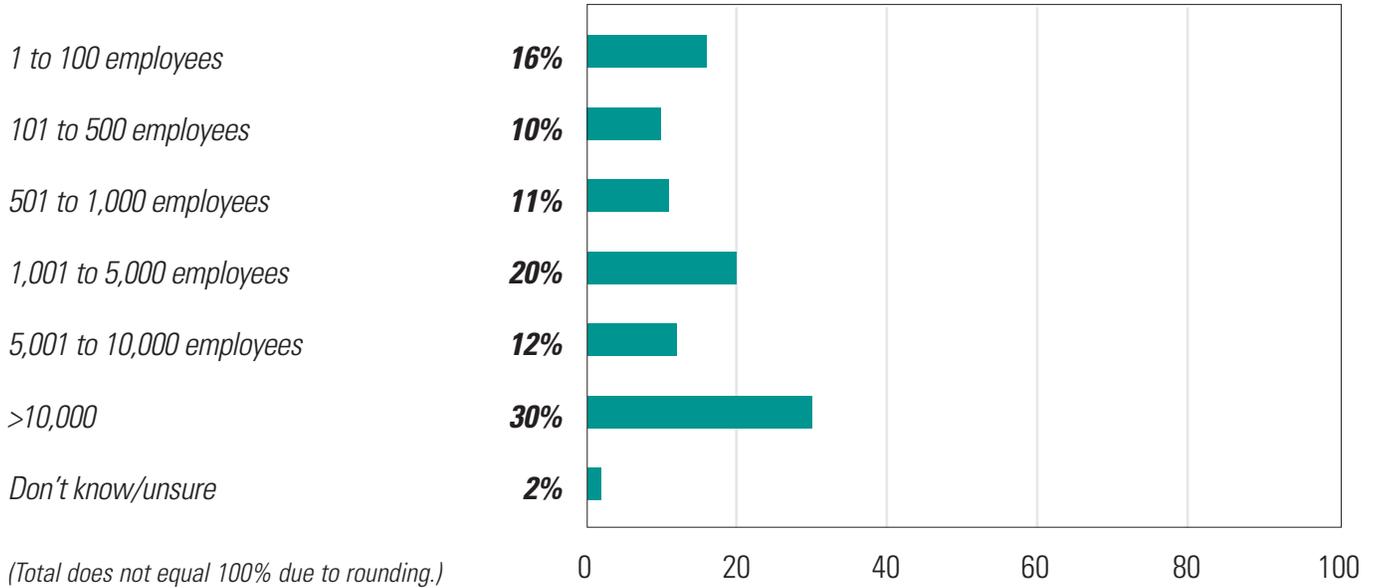
## DEMOGRAPHICS

### Figure 36: Respondents' Job Titles



## Figure 37: Number of Employees in Respondents' Organizations

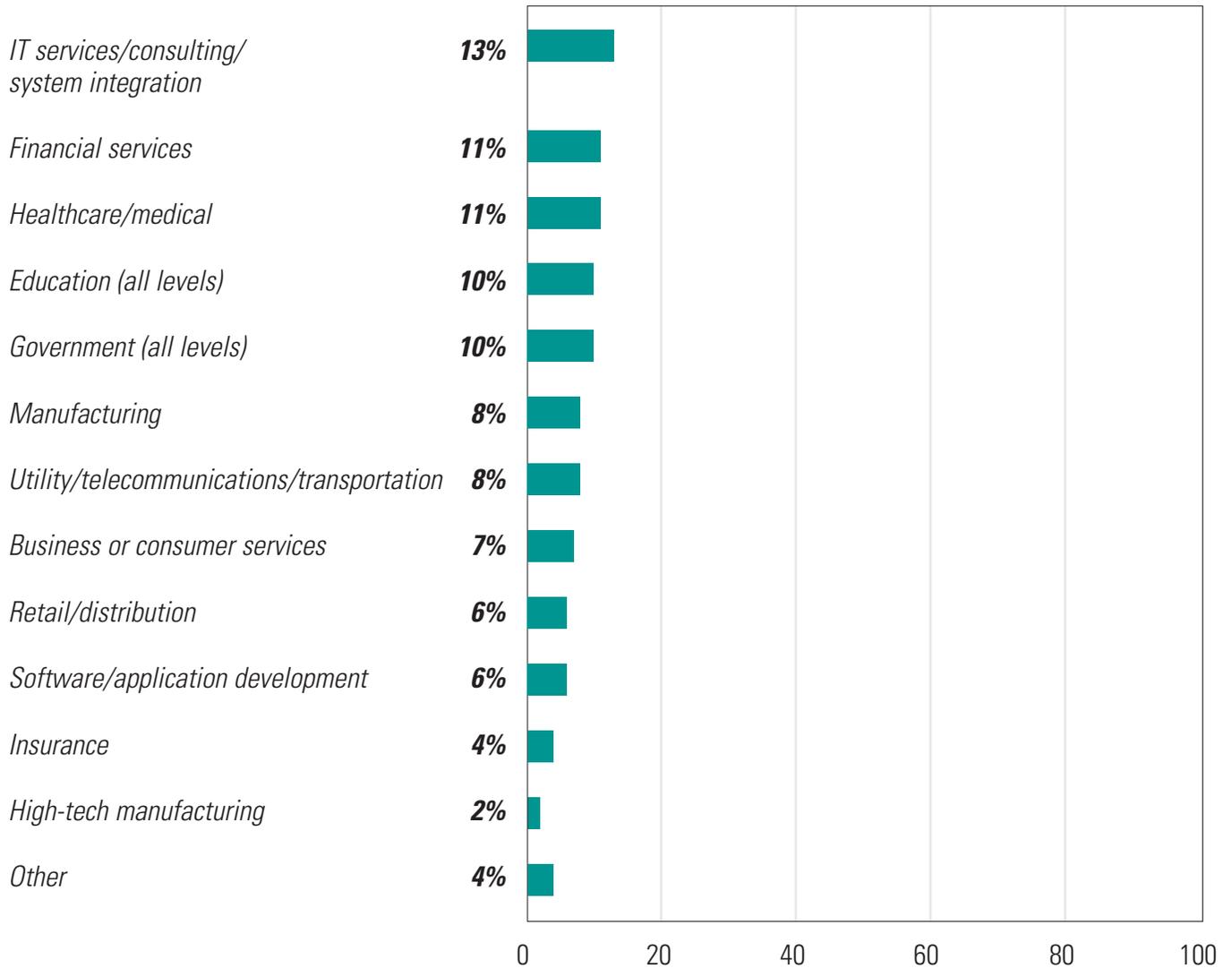
(Includes all locations, branches, and subsidiaries)



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.

## Figure 38: Respondents' Primary Industries



**Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey** was produced by Unisphere Research and sponsored by Oracle. Unisphere Research is the market research unit of Unisphere Media, a division of Information Today, Inc., publishers of Database Trends and Applications magazine and the 5 Minute Briefing newsletters. To review abstracts of our past reports, visit [www.dbta.com/About\\_Us#Unisphere](http://www.dbta.com/About_Us#Unisphere). Unisphere Media, 630 Central Avenue, Murray Hill, New Providence, NJ 07974; 908-795-3701, Email: [Tom@dbta.com](mailto:Tom@dbta.com), Web: [www.dbta.com](http://www.dbta.com).

Data collection and analysis performed with SurveyMethods.