



VENTANA RESEARCH



# Big Data Integration

Challenges and Opportunities in Accessing and  
Using Today's Information

White Paper



**Sponsors**

**May 2014**



Ventana Research performed this research to determine attitudes toward and utilization of big data integration. This document is based on our research and analysis of information provided by organizations that we deemed qualified to participate in this benchmark research.

This research was designed to investigate big data integration practices and needs and potential benefits. It is not intended for use outside of this context and does not imply that organizations are guaranteed success by relying on these results to improve the use of big data. Moreover, gaining the most benefit from big data integration requires an assessment of your organization's unique needs to identify gaps and priorities for improvement.

The full report with detailed analysis is available for purchase. We can provide detailed insights on this benchmark research and advice on its relevance through the Ventana On-Demand research and advisory service. Assessment Services based on this benchmark research also are available.

We certify that Ventana Research wrote and edited this report independently, that the analysis contained herein is a faithful representation of our evaluation based on our experience with and knowledge of big data and information management, and that the analysis and conclusions are entirely our own.

**Ventana Research**  
2603 Camino Ramon, Suite 200  
San Ramon, CA 94583-9137  
[info@ventanaresearch.com](mailto:info@ventanaresearch.com)  
(925) 242-2579  
[www.ventanaresearch.com](http://www.ventanaresearch.com)



# Table of Contents

## **Executive Summary 4**

### **Key Insights 10**

- Organizations are working to address big data integration. 10
- Data integration is demanding work that benefits from using dedicated technology. 11
  - Organizations often lack confidence and satisfaction in dealing with big data. 11
  - IT and business should work together on big data integration. 13
  - Big data requires specific capabilities for data integration. 13
- Big data integration must adapt to the prevalence of cloud computing systems. 14
  - Analytics is an essential complement to big data integration. 15
- Evaluating and selecting big data integration technology can be a challenge. 16
  - Organizations are uncertain about choosing big data integration technology. 16
  - Hadoop is a key technology for big data integration. 17
  - Big data integration can yield significant benefits. 18

### **10 Best Practice Recommendations 19**

- Evaluate the efficiency of your data integration processes. 19
  - Identify the business benefits of big data integration. 19
- Assess your organization's current approach to big data and integration. 20
  - Consider using dedicated technology for big data integration. 20
  - Determine the capabilities needed for big data integration. 20
- Establish collaboration between IT and business on big data integration. 21
  - Include cloud-based sources in your integration plans. 21
  - Consider the role of analytics in deriving value from big data. 22
  - Evaluate big data integration technology systematically. 22
- Decide which types of vendors and deployment methods best fit your needs. 23

### **About Ventana Research 24**

### **Appendix: About This Benchmark Research 25**

- Methodology 25
- Qualification 25
- Demographics 26
  - Company Size by Workforce 26
  - Company Size by Annual Revenue 27
  - Geographic Distribution 27
  - Industry 28
  - Job Title 28
- Role by Functional Area 29

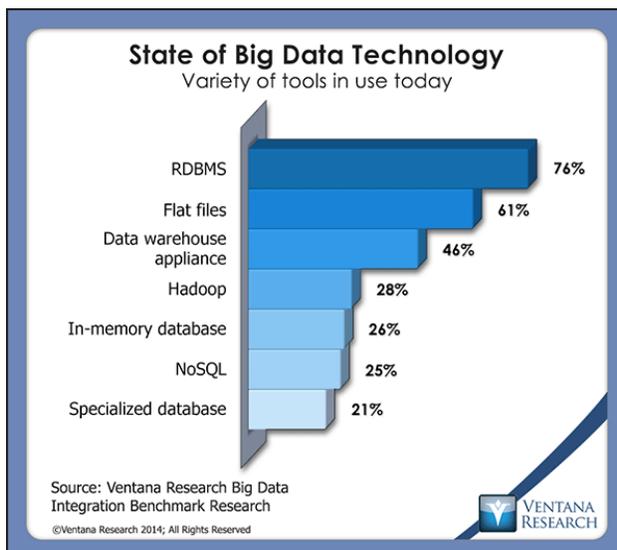


## Executive Summary

Data continues to flood into today's enterprises in ever-increasing velocity, variety and volume. This big data brings with it challenges – in storing it and in integrating it all into a form that can be used for business tasks. Many organizations try to use technology already on hand to collect, access and integrate big data. But processing manually or using legacy tools is slow and risks creating errors that undermine the value of the information and cause users to lose confidence in it. Automated processes using technology specifically designed for big data integration can overcome these issues and enable businesses to use the information to make decisions.

Ventana Research undertook this benchmark research to investigate accessing and transforming big data for processing and choosing technologies and tools to facilitate these efforts. We examined how organizations use big data integration or are planning to use it. We also assessed more general aspects of data integration and technologies that big data requires.

The use of big data techniques has become widespread: Almost half (48%) of all organizations participating in this research and two-thirds of the very large ones use it for storage, and 45 percent intend to use big data either in the next year or sometime in the future. The proliferation of data sources is a major driver for adoption of big data. More than one in five (22%) organizations now use more than 20



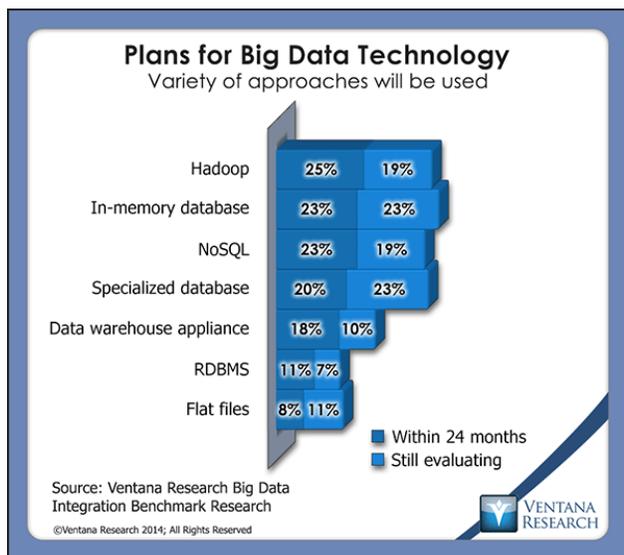
sources, and almost as many (19%) use 11 to 20 data sources. The sources of information most important to integrate with big data environments are applications (for 72%), customer data sources (57%) and data warehouses (56%).

Yet the research shows that only one in three (32%) participating organizations are satisfied with their current data integration technology. The most widely used are conventional: relational database management systems (RDBMSs, 76%) and



flat files (61%), which are not designed for this purpose. More than half (55%) of organizations use their existing data integration technology for big data integration as well, but only 29 percent of them are satisfied with it in this context; the same percentage are neutral or not satisfied. The main reasons cited for dissatisfaction are that the infrastructure is not fast enough and the technology is not adaptable to change (each by 56%); 44 percent said that the technology is simply not adequate.

The research also finds that more than half (55%) plan to change the way they assess and select big data integration technology in the next 12 to 18 months. Specifically, one in four said they will adopt Hadoop,



a specific type of big data technology, within 24 months, and about one-fourth plan to embrace other approaches in the next two years, including in-memory databases, NoSQL tools and specialized databases. Among the minority already using dedicated data integration technology, nearly all (93%) are satisfied with it. We regard these findings as evidence that appropriate tools are necessary to deal successfully with the demands of big data.

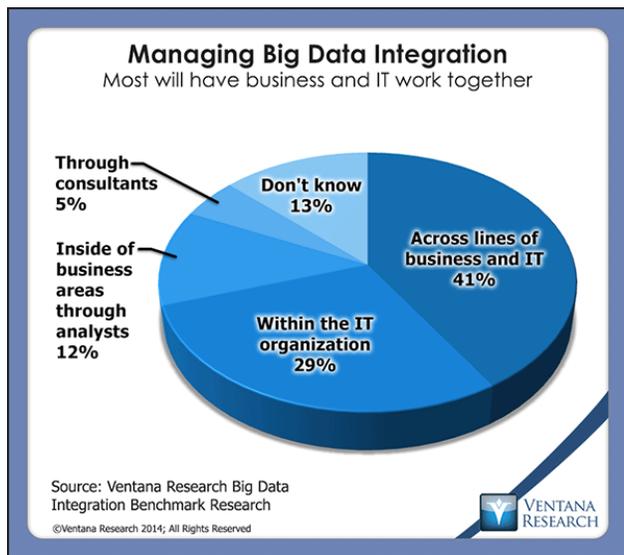
The research also finds that organizations waste time on data management tasks and that can impede the efficiency of big data processes and projects. People spend the largest amounts of time in the analytic process in reviewing data for quality and consistency issues (52%), preparing data for integration (46%) and connecting to data sources for integration (39%). All of these consume time that could be used for actual integration. Dedicated tools that automate the preparatory tasks can free time for those more valuable activities.

We connect these inefficiencies in processes and tools to participants' feelings about how well they deal with big data. More than half (55%) said they are only somewhat confident or not confident in their ability to process large volumes of data. Even more (58%) are somewhat



confident or not confident in their ability to process data that arrives at high velocity. And fewer than one-fifth (18%) are satisfied with how they manage the storage of big data. Among those that use dedicated data integration technology, however, 86 percent are satisfied with how they manage their storage.

The research shows that many organizations integrate big data technology with other technology in use in the business: Two out of three integrate it with business intelligence systems (67%), applica-



tions running in business processes (66%), data warehouses (65%) and business analytics managed in the lines of business (64%). This emphasis on business integration is reflected in how organizations plan to manage big data integration technology: 41 percent will do it across the lines of business and IT, and an even higher percentage of large organizations (59%) will do this. Only 29 percent will leave integration management entirely within the IT organization, and just 12 percent assign the task exclusively to

business analysts. Similarly, the largest percentage (44%) have business analysts work with IT to design and deploy big data integration; one-third assign IT alone to build the integration.

Nevertheless, IT is the area that uses big data and needs integration the most (55%), and half of organizations fund big data initiatives from the general IT budget; line-of-business IT budgets (38%) are the second-most commonly used. Despite the need for IT leadership in this technical area, we conclude that big data is critical to business users and thus recommend that managing it be a shared activity. We see this sentiment in the research also: More than four out of five (84%) participants said that collaboration in tasks related to big data integration is important or very important.

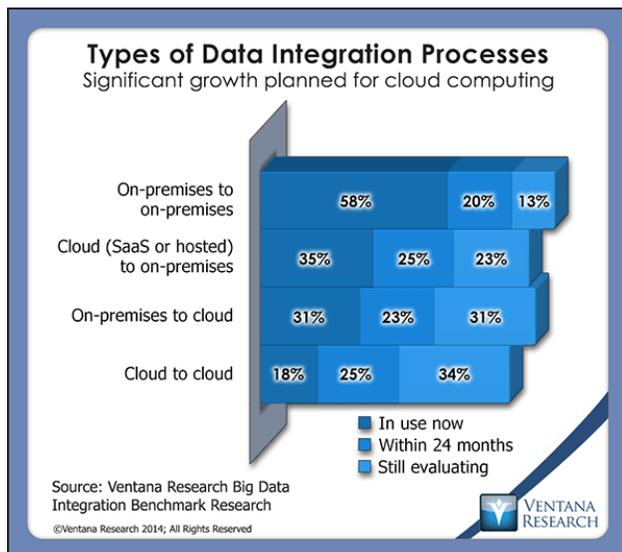
The two parties should work out how business users participate. The research shows that the data integration capabilities most often called critical for big data efforts are highly technical and thus primarily the



concern of IT: More than half named developing and managing metadata that can be shared across BI systems; joining disparate data sources during transformation; and establishing rules for processing and routing data. At the system level, conventional enterprise capabilities are most often needed: load balancing (cited by 51%), cross-platform support (47%), a development and testing environment (42%), systems management (40%) and scalable execution of tasks (39%). More likely to involve business users is the demand for mobile access to monitor big data integration, which nearly half (48%) of all participants said is important or very important.

Along with collaboration and mobility, two other innovative technology trends that we track have roles in big data integration. One is cloud computing, which half of organizations plan to use to manage big data. However, only 28 percent prefer to use it to deploy big data integration software, compared to almost half (46%) that prefer to integrate big data on-premises. This is not an unusual situation when

organizations deploy new types of software, especially those related to information management.



A more subtle analysis has to do with where data is stored and integrated. While organizations most commonly integrate data between on-premises systems (58%), one in three (35%) move cloud-based data on-premises, and nearly as many (31%) move data in the opposite direction.

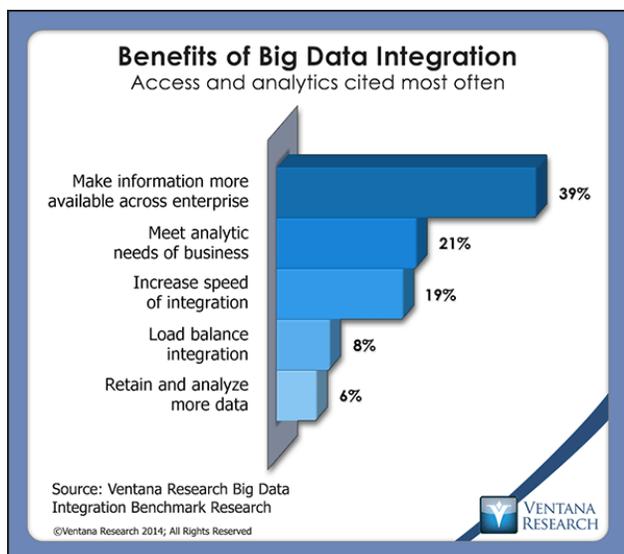
Looking ahead we find that over the next two years organizations plan data integration that involves cloud systems in some way more often than between on-premises systems. An indication of this as a cutting-edge trend is that more organizations already using cloud-to-cloud integration are confident in their abilities to process large volumes of big data and big data arriving at high velocity than are those that rely on on-premises systems.

The other innovative technology, analytics, is closely associated with big data as analytics are necessary to derive insights usable for



decision-making from big data. Two out of three organizations integrate big data technology with business intelligence systems and business analytics managed in the lines of business. Asked about types of analytics that require big data integration, more than half of organizations named data exploration (67%), predictive analytics (61%), query and reporting (60%), data discovery (58%) and forecasting, planning and what-if analysis (57%). This is a symbiotic relationship: Before analytics can provide value, the big data it operates on must be integrated and processed.

Processing and analyzing big data already is a proven solution for organizations being inundated by data and desperate to understand it. More than half of organizations (56%) in this research that use big data technology said they have improved their activities and processes in this regard. Almost all (93%) those that use a dedicated data integration product said they have improved their activities and



processes, as did more than four-fifths (84%) of those that use big data technology for storage. For big data integration specifically, the benefit that organizations most often ranked most valuable (39%) is that it makes information more available in a consistent manner; that it meets the analytics needs of business more easily was ranked first or second by half of participants.

Looking ahead, this research finds increasing interest in tools designed for big data integration. While only 12 percent of organizations have used dedicated technology for this purpose for more than a year or began using it in the last 12 months, three times as many (39%) said they will begin to use it within a year; the largest percentage (46%) said they intend to use such technology but do not know when they will do so. As organizations understand the challenges of big data integration we expect to see more use of dedicated tools.

Businesses should develop a comprehensive strategy for managing data and turning it into actionable information. Integration and



processing of big data, and subsequent use of analytics on the data, are essential steps in doing so. Capable technology tools can make it smoother and faster and help produce better guidance toward achieving their goals.



## Key Insights

This benchmark research yielded the following important general findings and key insights regarding the application of data integration techniques to big data to meet IT and business needs. (We discuss performance levels in the Performance Index portion of the full research report; the actual questions asked in our survey are in an appendix to the research report. Specifics of organization sizes are in the appendix "About This Benchmark Research.")

### **Organizations are working to address big data integration.**

The findings of this research make clear that organizations are aware of the need to integrate big data and are beginning to address it. Our

The findings of this research make clear that organizations are aware of the need to integrate big data and are beginning to address it.

Performance Index analysis finds an almost even distribution across the lower three of the four levels into which we segment organizations' performance in big data integration, with 27 to 33 percent at each of those levels. At the top end, however, only 13 percent reach the highest Innovative level of performance. Analyzed by industry, manufacturing is the most mature sector, with more than one-fifth (22%) performing at the Innovative level.

Measured by both number of employees and annual revenue, very large companies (only 23% of all participants) have the largest percentage at the Innovative level (22% and 24% respectively); these organizations usually have access to enough personnel and financial resources to perform well in big data integration.

More detailed analysis shows that of the four dimensions that comprise performance, companies perform best in the Information dimension of big data integration: Almost one-fifth (17%) are Innovative and almost one-third (31%) are Strategic. Organizations perform least well in the People dimension, as our analysis places two-thirds at the two lowest levels. One reason for the low level of People performance is spotty training in the skills required to integrate big data. In several areas, almost as many participants said the training they provide is only somewhat adequate or inadequate as said it is mostly or



completely adequate. Overall, our analysis of the research places more than half of participants in the lower half of the performance hierarchy for all four dimensions.

### **Data integration is demanding work that benefits from using dedicated technology.**

For most IT organizations data integration has become an essential activity, for a simple reason: To stay competitive, the enterprise must be able to process ever-increasing amounts of data in ever more forms

Only one in three (32%) of all those participating in this research are satisfied with their current data integration technology.

from an increasing number of sources. But most have difficulty in attaining that goal: Only one in three (32%) of all those participating in this research are satisfied with their current data integration technology.

One-third (34%) of them are using a single vendor for data integration. That's significant, as analysis shows that nearly all organizations (93%) using dedicated data integration technology are satisfied with it.

Moreover, four out of five (80%) of them are satisfied with their data integration technology's support for big data, compared to 56 percent of those that do not use a single vendor for data integration. The research thus finds that organizations using a dedicated approach are significantly more satisfied with their efforts.

These findings explain the increasing interest in tools purpose-built for big data integration. While only 12 percent have used dedicated technology for this purpose for more than a year or began using it in the last 12 months, more than three times as many (39%) said they will begin to use it within a year; the largest percentage (46%) said they intend to use such technology but do not know when they will do so. As organizations understand the challenges of big data integration we expect to see more use of dedicated tools.

These findings explain the increasing interest in tools purpose-built for big data integration. While only 12 percent have used dedicated technology for this purpose for more than a year or began using it in the last 12 months, more than three times as many (39%) said they will begin to use it within a year; the largest percentage (46%) said they intend to use such technology but do not know when they will do so. As organizations understand the challenges of big data integration we expect to see more use of dedicated tools.

### **Organizations often lack confidence and satisfaction in dealing with big data.**

Big data is still new to many organizations, and the research finds that many are concerned about processing large volumes of data. More than half (55%) of participating organizations characterized



themselves as either somewhat confident or not confident in their ability to accomplish that task. They have even less confidence in their ability to process data that arrives at high velocity: 29 percent each said they are somewhat confident or not confident. When it comes to integrating the variety of big data the company needs to perform well, confidence is somewhat stronger, with more than half (56%) declaring themselves confident or very confident. Not surprisingly, 86 percent of the organizations that are very confident in their ability to integrate the variety of big data they need also are satisfied with how they manage the storage of big data. We find similar levels of confidence in the data quality of the information being generated: More than half (54%) are confident or very confident in it. Likewise nearly all (91%) of the organizations that are very confident in that data quality are satisfied with the way they manage storage of big data.

But looked at overall, there is considerably less satisfaction with organizations' management of their big data storage: Fewer than one-fifth (18%) of all participants said they are satisfied with how they manage their storage; most (58%) said they are somewhat satisfied, which indicates room for improvement. However, many more organizations that use dedicated data integration technology (86%) are satisfied than those that don't use dedicated technology (52%).

Making a commitment to data integration, for big data and otherwise, can pay off in confidence and satisfaction with the processes.



Only one-third (32%) of organizations are satisfied with their current data integration technology, while twice as many (66%) are satisfied with their data integration processes for loading and creating big data. A substantial majority (86%) of those very confident in their ability to integrate the needed variety of big data are satisfied with their existing data integration processes.

Those that are not satisfied said the process is too slow (61%), analytics are hard to build and maintain (50%) and data is not readily available (39%). In our view, these findings indicate that making a commitment to data integration, for big data and otherwise, can pay off in confidence and satisfaction with the processes.



### **IT and business should work together on big data integration.**

The largest percentage of participants in this research (44%) use business analysts working with IT to design and deploy big data integration. Another one-third assign IT to build the integration, and half that percentage (16%) have IT use a dedicated data integration tool. The research finds some resistance to having business users integrate big data that IT has not prepared first: Almost one in four participants (23%) said they are resistant or very resistant to that, and the majority (51%) resist somewhat; only 27 percent said they are not resistant. For more than half (58%) the IT group responsible for BI and data warehouse systems also is the key stakeholder for designing and deploying big data integration; no other option is used by more than 11 percent, as only 6 percent each use a centralized data management team or an IT function within lines of business. Along the same lines, IT is the area that uses big data and needs integration the most (55%). The most frequent issue arising between business units and IT is entrenchment of budgets and priorities (in 42% of organizations). Funding of big data initiatives most often comes from the general IT budget (50%); line-of-business IT budgets (38%) are the second-most commonly used. It is understandable that IT dominates this heavily technical function, but big data is beneficial only when it advances the organization's goals for information that is needed by business. Management should ensure that IT works with the lines of business to enable them to get the information they need to improve business processes and decision-making.

Management should ensure that IT works with the lines of business to enable them to get the information they need to improve business processes and decision-making.



### **Big data requires specific capabilities for data integration.**

The data integration capabilities that most participants said are critical for their big data efforts are to develop and manage metadata that can be shared across BI systems (cited by 58%), to join disparate data sources during transformation (56%) and to establish rules for processing and routing data (56%). For key big data integration tasks more than half of participants said their processes are mostly or completely adequate: accessing (63%), loading (60%), extracting



(59%), archiving (55%) and copying (52%). On the other side of the coin, the tasks most often identified as only somewhat adequate or inadequate are quality (46%), analytics (45%) and auditing (44%). We note that these and other capabilities are found in data integration products available today.

At the system level, the research finds that conventional enterprise capabilities are most often needed: load balancing (51%), cross-platform support (47%), a development and testing environment (42%), systems management (40%) and scalable execution of tasks (39%). To test the range of big data integration capabilities before it is applied to production projects, the “sandbox” has become the standard approach. For their development and testing environment, the largest percentage (36%) will have an internal sandbox with specialized big

Among innovative technology trends that we track, this research finds that collaboration and mobility are gaining momentum.

data, while almost one in five (17%) will have an internal sandbox without specialized big data. Fewer organizations will have a hybrid combination of internal and external sandboxes (18%), and only 10 percent will use an external one.

Among innovative technology trends that we track, this research finds that collaboration and mobility are gaining momentum. Collaboration in tasks related

to big data integration is important (40%) or very important (44%) for many organizations, and mobile access to monitor big data integration is important or very important to nearly half (48%). Many of these integration priorities can contribute to the success of big data processes and technologies.

### **Big data integration must adapt to the prevalence of cloud computing systems.**

Predictably, the research finds that data integration processes are most often applied between systems deployed on-premises (58%). Second-most prevalent is integrating cloud-based systems with those on-premises (35%), which reflects the progress cloud computing has made. Nonetheless, cloud-to-cloud integration remains least common (18%). In the next year or two 20 to 25 percent of organizations plan additional support for all types of integration; those being considered the most are cloud-to-cloud (25%) and on-premises-to-cloud (23%).



Organizations already using cloud-to-cloud integration are more often confident in their abilities to process large volumes of big data (20%) and to process big data arriving at high velocity (14%) than are other organizations. In addition, nearly all (95%) organizations using cloud-to-cloud integration said they have improved their activities and processes. This finding confirms the value of integration of big data regardless of what types of systems hold it. Those that are doing point-to-point integration like that in cloud computing are more confident than others. With a growing number of organizations using cloud computing, data integration is a critical requirement for big data projects.

With a growing number of organizations using cloud computing, data integration is a critical requirement for big data projects.



### **Analytics is an essential complement to big data integration.**

Analytics and business intelligence are technological companions for big data. Two out of three organizations integrate big data technology with business intelligence systems (67%), applications running in business processes (66%), data warehouses (65%) and business analytics managed in the lines of business (64%). Analytics can add value to big data investments by providing new insights for decision-makers. In this research half of participants ranked analytics first among six innovative technologies that can improve the integration of big data. Analytics also is the data activity participants said is most often required (by 87%) for integrating big data, followed by quality of data (68%), extraction (63%) and security (61%). The research finds that among types of analytics that require big data integration, more than half of organizations named data exploration (67%), predictive analytics (61%), query and reporting (60%), data discovery (58%) and forecasting, planning and what-if analysis (57%). More than two in five cited information discovery (49%), real-time analytics (46%) and search (also 46%). Organizations that intend to apply analytics to big data must first ensure that the data is integrated and prepared for analysis.



## Evaluating and selecting big data integration technology can be a challenge.

For a process as complex as big data integration, choosing the right technology tool can be difficult. The research shows that the most important technology and vendor criteria applied to evaluate participants' current data integration software are usability (important for 80%), reliability (73%) and functionality (66%). The same three top the list in prepurchase evaluation of big data integration technology: The largest percentages of very important criteria are for usability (for 53%), reliability (52%) and functionality (49%).

To judge reliability, participants most often said they look at the IT architecture (63%) and data performance (60%). For manageability (which 42% identified as very important), IT administration (78%) is by far the aspect considered most. Slightly more than half (52%) said they are satisfied with the results produced by the group that is identifying and selecting the organization's big data integration

More than half (55%) of organizations are planning to change the way they assess and select big data integration technology.

technology. Nearly one-third had no opinion about this topic, but for the 16 percent who declared themselves not satisfied, the most common reason is that the group takes too long (53%); nearly as many (47%) said the group doesn't communicate its strategy.

Despite general agreement about these issues among research participants, more than half (55%) of organizations are planning to change the way they assess and

select big data integration technology. The top three factors driving that change are business improvement initiatives (for 54%), wider analytics and BI initiatives (46%) and improvement in the quality of business processes (40%). Thus the research finds that big data integration, which can serve many needs in an organization, is still in flux in terms of processes and tools.

## Organizations are uncertain about choosing big data integration technology.

Evaluations of big data integration technology also should include consideration of how to deploy it and what sort of vendors can provide



it. Almost half (46%) of organizations prefer to integrate big data on-premises while 28 percent opt for cloud-based software as a service and 17 percent expressed no preference. At the same time, half of organizations plan to use cloud computing for managing big data; another one-third (32%) don't know whether they will. Among the less than one-fifth (18%) that explicitly said they will not use the cloud, the overwhelmingly most frequent reason is security (84%); this is a longstanding bias regarding cloud computing and one that we find increasingly less viable.

When it comes to vendors being considered to supply big data integration technology, other than the current data management vendor the most-often named are the current database (51%) and data integration (49%) vendors. Only 38 percent would look to a specialized vendor. The most common barrier to improving big data integration is cost of the software or license (44%), followed by a lack of resources to use on improvement (37%) and the sense that big data technologies are too complicated to integrate (35%).

Along with identifying solid business benefits, establishing savings of time and money are essential pieces of a convincing rationale for investment in big data integration technology.

Along with identifying solid business benefits, establishing savings of time and money are essential pieces of a convincing rationale for investment in big data integration technology.

### **Hadoop is a key technology for big data integration.**

The range of big data technologies continues to expand to meet demands for the storage and use of information across the enterprise. An early and fast-growing technology is the open source Apache Hadoop and commercial enterprise versions of it; Hadoop has a distributed file system that is used to manage and interface to other systems. Currently 28 percent of organizations use Hadoop, but the research shows that more organizations (25%) plan to use it than other technologies in the next two years. Hadoop-specific skills to support big data integration are found in 47 percent of organizations today. For those that are resource-limited open source Hadoop can be affordable, and to automate and interface with it adopters can use SQL



in addition to its native interfaces; about three in five organizations now use each of these options. Hadoop can be a significant means to implement big data and should be integrated with other systems.

### **Big data integration can yield significant benefits.**

To gain approval and funding for any technology investment, proponents typically must demonstrate significant business benefits.

Almost all (93%) those that use a dedicated data integration product said they have improved their activities and processes by using it.

More than half of organizations (56%) in the research that use big data technology said they have improved their activities and processes significantly (14%) or slightly (42%) by using it. Almost all (93%) those that use a dedicated data integration product said they have improved their activities and processes, as did more than four-fifths (84%) that use big data technology for storage. For big data integration specifically organizations most often ranked first as a realized benefit that it makes information more available in a consistent manner (39%); that it meets the analytics needs of business more easily was

ranked first (21%) or second (29%) by half of participants. These also are the benefits most often expected by those planning to use big data integration. The organizational areas most often benefitting from big data are IT (57%), customer service (44%), operations (43%) and marketing (40%). Those planning a big data initiative and use dedicated data integration expect to have significant benefits.



## 10 Best Practice Recommendations

This benchmark research reveals significant new insights into the evolving nature and use of big data integration processes and systems. For organizations considering how to optimize their use of big data, we offer the following recommendations.

### **1. Evaluate the efficiency of your data integration processes.**

The research finds a significant amount of time wasted on data management tasks that can impede the efficiency of big data processes and projects. The largest amounts of time in the analytic process are spent in reviewing data for quality and consistency issues (in 52% of organizations), preparing data for integration (46%) and connecting to data sources for integration (39%). All of these consume more time than is available for actual integration tasks. Look for software that automate these preparatory tasks. Also take into account the frequency with which the organization needs to integrate data into big data stores: More than one-third (37%) do it hourly or daily, and one in four others need to do it in real time. Require that any system you evaluate be able to do this quickly enough and be flexible enough to recycle your big data at different frequencies based on business need, which almost half (42%) of participants do.

### **2. Identify the business benefits of big data integration.**

As in the case of any other proposed technology investment, early in the process identify the most significant business benefits you seek from big data integration. More than half of organizations (56%) that use big data technology said they have improved their activities and processes by using it. Identify the ones most important for your initiative. For big data integration specifically, both organizations that already use it and others that intend to most often ranked first the benefit of making information more available in a consistent manner, followed by meeting the analytics needs of the business. Also identify the areas of the organization that will benefit most from using big data; for those participating in this research they are IT (57%), customer service (44%), operations (43%) and marketing (40%). Involve knowledgeable stakeholders from each area in planning and evaluation.



### **3. Assess your organization's current approach to big data and integration.**

The research finds that many organizations are concerned about processing big data and the speed at which they receive it. More than half (55%) said they are only somewhat confident or not confident in their ability to process large volumes of data. Even more (58%) are somewhat confident or not confident in their ability to process data that arrives at high velocity. Fewer than one-fifth (18%) of all participants are satisfied with how they manage the storage of big data. Related to these anxieties is the finding that only one-third (32%) are satisfied with their current data integration technology. Evaluate all these aspects of your big data strategy, as well as the ability to integrate the variety of big data, and make changes to improve deficiencies.

### **4. Consider using dedicated technology for big data integration.**

Only one-third of all participants in this research are satisfied with their current data integration technology. Compare the effectiveness of your tools with that of others designed for this purpose. Almost all (93%) organizations that use a dedicated data integration product said they have improved their activities and processes, as did more than four-fifths (84%) that use big data technology for storage. In addition, four out of five (80%) of them are satisfied with their data integration technology's support for big data. At this time only 12 percent use dedicated technology for this big data integration, but we find a clear trend in that direction: 39 percent said they will begin to use it within a year, and even more (46%) intend to use such technology but do not know when they will do so. Acting soon to acquire the right tools could give you a competitive advantage in using big data.

### **5. Determine the capabilities needed for big data integration.**

The data integration capabilities that most participants said are critical for their big data efforts are to develop and manage metadata that can be shared across BI systems (cited by 58%), to join disparate data sources during transformation (56%) and to establish



rules for processing and routing data (56%). At the system level, conventional enterprise capabilities are most often needed: load balancing (51%), cross-platform support (47%), a development and testing environment (42%), systems management (40%) and scalable execution of tasks (39%). Identify the most important for your organization and look for them in products you consider. Also test the adequacy of the full range of integration tasks; in descending order 63 to 52 percent of research participants said these are mostly or completely adequate: accessing, loading, extracting, archiving and copying. The tasks most often identified as only somewhat adequate or inadequate are quality (46%), analytics (45%) and auditing (44%). Make sure you can cover the basics and then seek more advanced capabilities.

## **6. Establish collaboration between IT and business on big data integration.**

Collaboration in tasks related to big data integration is important (40%) or very important (44%) to most organizations. We believe this should apply to the relations between IT staff and business users in this basically technical aspect of information management, even though IT is the area that uses big data and needs integration the most (55%). We note that only about one in four organizations are not resistant to having business users integrate big data that IT has not prepared first. However, the largest percentage (44%) have business analysts work with IT to design and deploy big data integration, and big data is beneficial only when it feeds the business's need for information. Take steps to ensure that both sides understand what is at stake and build a process in which all stakeholders participate and benefit as is appropriate.

## **7. Include cloud-based sources in your integration plans.**

Although the majority (58%) of data integration processes are applied between systems deployed on-premises, one-third (35%) of organizations are integrating cloud-based systems with those on-premises. In the next year or two about one in four will add support for the other end points of integration: from cloud to cloud (25%) and on-premises to cloud (23%). Assess your organization's current cloud deployments and foreseeable use of others, then include them in your planning. Big data has value regardless of what types of



systems hold it, so we advise taking an expansive view of integrating it.

## **8** Consider the role of analytics in deriving value from big data.

Most organizations apply analytics and business intelligence to acquire insights for decision-makers from their big data. Half of all participants ranked analytics first among six innovative technologies that can improve the integration of big data. Two out of three integrate big data technology with business intelligence systems, applications running in business processes, data warehouses and business analytics managed in the lines of business. Participants also cited a variety of useful analytics that require big data integration, led by data exploration (67%), predictive analytics (61%), query and reporting (60%), data discovery (58%) and forecasting, planning and what-if analysis (57%). Understanding the mass of data they receive is a major impetus for organizations undertaking big data initiatives, and analytics is a key tool for doing that. Include it in thinking about integration.

## **9** Evaluate big data integration technology systematically.

For a process as complex as big data integration, choosing the right technology tool can be difficult. The research shows that the most important technology and vendor evaluation criteria for both participants' current data integration software and their prepurchase evaluation of big data integration technology are usability, reliability and functionality. These criteria likely will be important to the more than half (55%) of organizations planning to change the way they assess and select big data integration technology. The top three factors driving that change are business improvement initiatives (for 54%), wider analytics and BI initiatives (46%) and improvement in the quality of business processes (40%). In your own evaluations, combine the qualities desired in products with their suitability for the kinds of projects you intend to use big data for.



## **10. Decide which types of vendors and deployment methods best fit your needs.**

■ It is logical to look to one's current data management vendor for big data integration technology, assuming that the organization is satisfied with its product and services. Other than that source, most participants said they would consider their current database (51%) and data integration (49%) vendors. Only 38 percent would look to a specialized vendor. Big data is still rather new, so don't assume that suppliers you are comfortable with are on the leading edge here; make them prove it. Unless you are wholly satisfied with that proof, broaden your search at least to build a basis for comparison. Of course, include expense in your analysis; cost of the software or license (cited by 44%) is the most common barrier to improving big data integration. More than one-third (35%) said the sense that big data technologies are too complicated to integrate is a barrier. Assure yourselves that prospective suppliers can overcome it.



## About Ventana Research

Ventana Research is the most authoritative and respected benchmark business technology research and advisory services firm. We provide insight and expert guidance on mainstream and disruptive technologies through a unique set of research-based offerings including benchmark research and technology evaluation assessments, education workshops and our research and advisory services, Ventana On-Demand. Our unparalleled understanding of the role of technology in optimizing business processes and performance and our best practices guidance are rooted in our rigorous research-based benchmarking of people, processes, information and technology across business and IT functions in every industry. This benchmark research plus our market coverage and in-depth knowledge of hundreds of technology providers means we can deliver education and expertise to our clients to increase the value they derive from technology investments while reducing time, cost and risk.

Ventana Research provides the most comprehensive analyst and research coverage in the industry; business and IT professionals worldwide are members of our community and benefit from Ventana Research's insights, as do highly regarded media and association partners around the globe. Our views and analyses are distributed daily through blogs and social media channels including [Twitter](#), [Facebook](#), [LinkedIn](#) and [Google+](#).

To learn how Ventana Research advances the maturity of organizations' use of information and technology through benchmark research, education and advisory services, visit [www.ventanaresearch.com](http://www.ventanaresearch.com).



## Appendix: About This Benchmark Research

### Methodology

Ventana Research conducted this benchmark research on the Web from November 2013 through February 2014. We solicited survey participation via email, our website and social media invitations. Email invitations were also sent by our media partners and by vendor sponsors.

We presented this explanation of the topic to participants prior to their entry into the survey:

The scale of big data – the unrelentingly rapid rate at which it is generated as well as the size and types of data – adds to the challenge of getting it into a form that can be processed into big data stores. Without some form of automation to integrate the mass of big data, organizations risk delays and wasted time in processing it. This benchmark research explores the need for and benefits of accessing and integrating big data for processing.

The following promotion incited participants to complete the survey:

What's In It For You? Upon completion of the research, all qualified participants will receive a report on the findings of this benchmark research to support their organization's efforts, along with a \$5 Amazon.com gift certificate. In addition, all qualified participants will be entered into a drawing to win one of 25 benchmark research reports and a 30-minute consultation, a package valued at US\$1,495 or €1,232. Thank you for your participation!

### Qualification

We designed the research to assess the use of and plans for spreadsheets across organizations and industries. Qualification to participate was presented to participants as follows:

The survey for this benchmark research is designed for business and IT managers responsible for making big data ready to use or involved with the purchasing of technology for this area. Solution providers, software vendors, consultants, media and systems integrators may



participate in the survey, but they are not eligible for incentives and their input will be used only if they meet the qualifications. Incentives are provided to qualified participants in the research and also are conditional on provision of accurate contact information including company name and company email address that can be used for fulfillment of incentives.

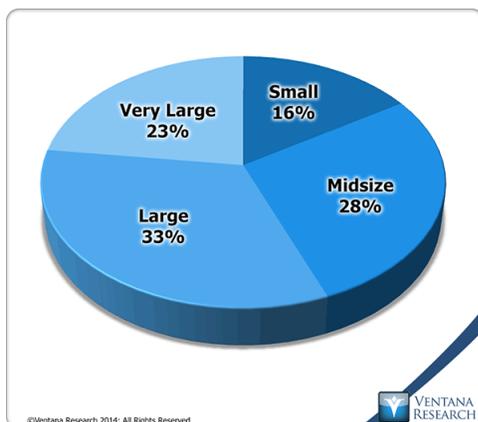
Further qualification evaluation of respondents was conducted as part of the research methodology and quality assurance processes. It entailed screening out responses from companies that are too small, questionnaires that were not materially complete, or those where the submission is from an inappropriate submitter or appears to be spurious.

## Demographics

We designed the survey used for this research to be answered by executives and managers across a broad range of roles and titles working in organizations. We deemed 216 of those who clicked through to this survey to be qualified to have their answers analyzed in this research. In this report, the term “participants” refers to that group, and the charts in this section characterize various aspects of their demographics and qualifications.

### Company Size by Workforce

We require participants to indicate the size of their entire company.



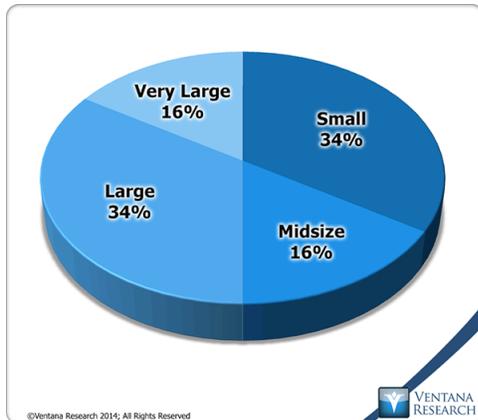
Our research repeatedly shows that size of organization, measured in this instance by employees, is a useful means of segmenting companies because it correlates with the complexity of processes, communications and organizational structure as well as the complexity of the IT infrastructure. In this research, participants represented a broad range of organization sizes in nearly equal numbers: 23 percent work in very large companies (having 10,000 or more employees), 33 percent work in large companies (with 1,000 to 9,999 employees), 28 percent work in midsize companies (with 100 to 999 employees), and 16 percent work in small companies (with fewer than 100 employees). This distribution is consistent with



prior benchmark research and our research objectives and provides a suitably large sample from each size category.

### Company Size by Annual Revenue

When we measured size by annual revenue, the distribution of



categories shifted downward; fewer companies fell into the very large and midsize categories and more than twice as many are small. By this measure, 7 percent fewer are very large companies (having revenue of more than US\$10 billion), 1 percent more are large companies (having revenue from US\$500 million to US\$10 billion), 12 percent fewer are midsize companies (having revenue from US\$100 to US\$500 million), and 18 percent more are small companies (with revenue of less than US\$100 million).

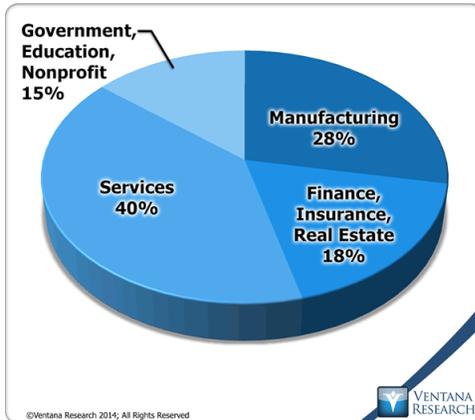
This sort of redistribution is typical in our research projects when we measure by revenue instead of head count.



### Geographic Distribution

A large majority (89%) of the participants were from companies located or headquartered in North America. Those based in Europe and in Asia Pacific accounted for 4 percent each, and those in the rest of the world accounted for 3 percent. This result was in keeping with our expectations at the start of this investigation, since organizations participating in our research most often are headquartered in North America. However, many

of these are global organizations operating worldwide.

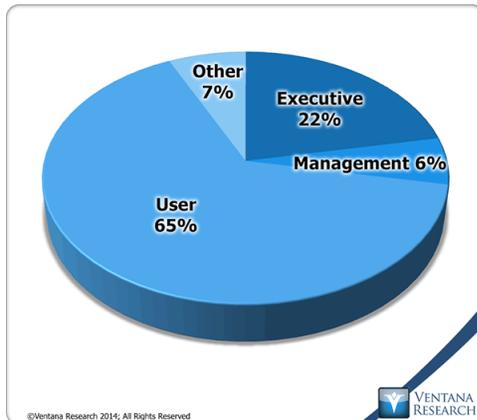


### Industry

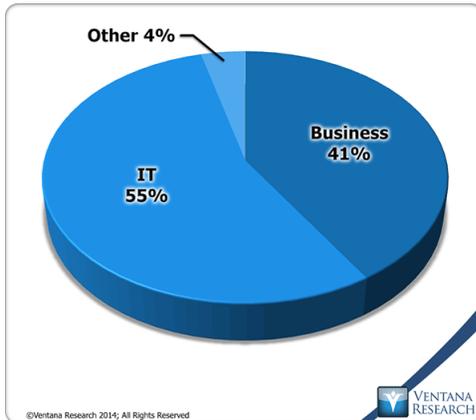
The companies of the participants in this benchmark research represented a broad range of industries, which we have categorized into four general categories as shown below. Companies that provide services and those in manufacturing accounted for two-thirds of the participants. Those in finance, insurance and real estate accounted for 18 percent, and those in government, education and nonprofits accounted for the balance.

### Job Title

We asked participants to choose from among 14 titles the one that best describes theirs.



We sorted these responses into four categories: executives, management, users and others. Slightly fewer than two-thirds identified themselves as having titles that we categorize as users, a grouping that includes director (25%), senior manager or manager (21%), analyst (10%) and staff (9%). A bit more than one-fourth are executives; most of them (15%) are CIOs. Another 6 percent are management, by which we mean vice presidents. Others, in this case consultants, accounted for the balance. We concluded after analysis that this response set provided a meaningfully broad distribution of job titles.



### Role by Functional Area

We asked participants to identify their functional area of responsibility as well. This enabled us to identify differences between participants who have differing roles in the organization. In this rather technical area more than half of the participants identified themselves as being in the IT function. The lines of business most often represented were operations (7%), marketing (6%) and executives or management (5%). Here again consultants made up the Other category.