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.
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%), applications 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.
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.
Appendix: About This Benchmark Research

Ventana Research designed this benchmark research for business and IT managers responsible for making big data ready to use or involved with the purchasing of technology for this area. The research was conducted from November 2013 through February 2014. Applying our standard methodology and quality assurance criteria, we identified 216 qualified participants. They represent a range of organization sizes: 23 percent from very large companies (having 10,000 or more employees), 33 percent from large companies (with 1,000 to 9,999 employees), 28 percent from midsize companies (with 100 to 999 employees), and 16 percent from small companies (with fewer than 100 employees). A large majority (89%) of these companies are located or headquartered in North America, although many of these are global organizations operating worldwide. Among industry categories, companies that provide services accounted for 40 percent, those in manufacturing for 28 percent and those in finance, insurance and real estate for 18 percent. Government, education and nonprofits accounted for the remaining 15 percent. Categorized by their job title, 22 percent are executives, 6 percent are in management, and the majority (65%) are what we term users in the lines of business. By functional area, 55 percent work in IT and 41 percent in business departments, most often operations (7%) and marketing (6%). (More demographic detail about the participants is available in the full research report.)

This Executive Summary is drawn from the full Ventana Research Benchmark Research report. The full report is available for purchase, payable by check or credit card. Advice and focused guidance based on this benchmark research can be purchased through our Ventana On-Demand service. For more information about the full Benchmark Research report or assessment of your organization using our Maturity Index methodology, please contact us at sales@ventanaresearch.com.