Welcome, Philip | My Account | Log Out

White Papers | Web Seminars | Newsletters | eBooks

- Big Data & Analytics
- Data Management
- MDM & Data Governance
- Infrastructure
- Info Strategy & Leadership
- BI & Data Discovery
- Mobility
- web seminars &
  white papers
- resource
  center

Blog

Big Data Requires Integration Technology

by Ventana Research
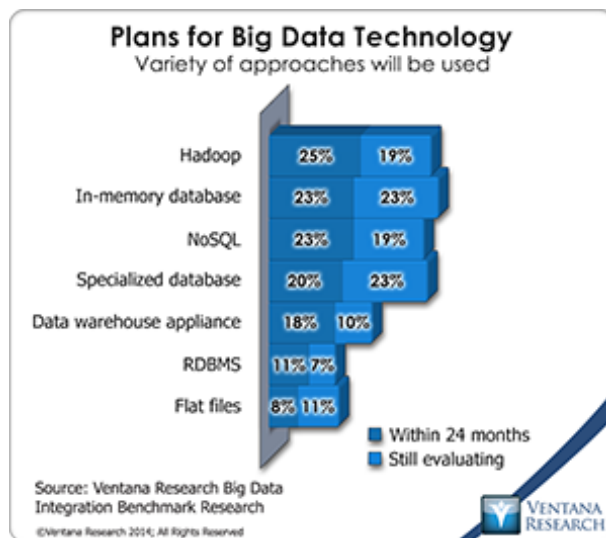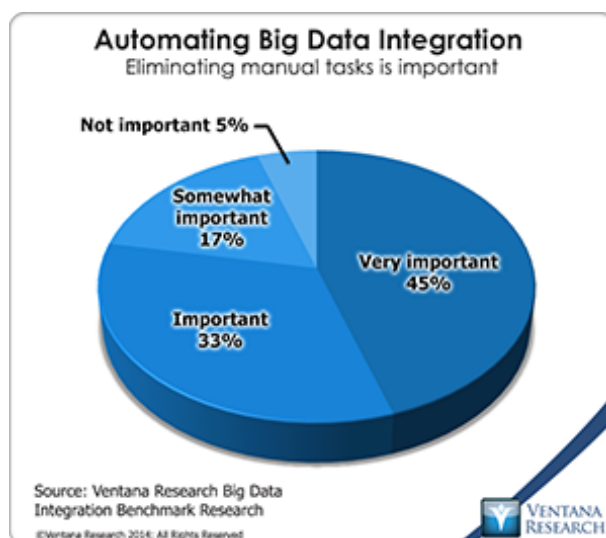DEC 18, 2014 8:48pm ET
Print
Email
Reprints
Comment
Twitter
LinkedIn
Facebook
Google+

The market for big data continues to grow as organizations try to extract business value from their own masses of data and other sources. Earlier this year I outlined the dynamics of the business opportunity for big data and information optimization. We continue to see advances as big data and associated information technologies

deliver more value, but the range of innovation also has created fragmentation among existing systems including databases that are managed onpremises or in cloud computing environments. In this changing environment organizations encounter new challenges not only in adapting to technology that is more efficient in automating data processing but also in integrating it into their enterprise architecture. I've already explained how big data can be ineffective without integration, and we conducted more in-depth research into the market, resulting in our benchmark research on big data integration, which reveals the state of how organizations are adopting this technology in their processes.

Plans for Big Data Technology
Variety of approaches will be used

Source: Ventana Research Big Data
Integration Benchmark Research
©Ventana Research 2014; All Rights Reserved

The research shows that 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 in the next year or sometime in the future. This is a significant change in that most organizations have used relational database management systems (RDBMSs) for nearly everything. We find that RDBMSs (76%) are still the most widely used big data technology, followed by flat files (61%) and data warehouse appliances (46%). But this is not the direction many companies are planning to take in the future: Hadoop (44%), in-memory database (46%), specialized databases (43%) and NoSQL (42%) are the tools most often planned to be used by 2016 or being evaluated. Clearly there is a revolution in approaches to storing and using data, and that introduces both opportunities and challenges.

Automating Big Data Integration
Eliminating manual tasks is important

Source: Ventana Research Big Data
Integration Benchmark Research
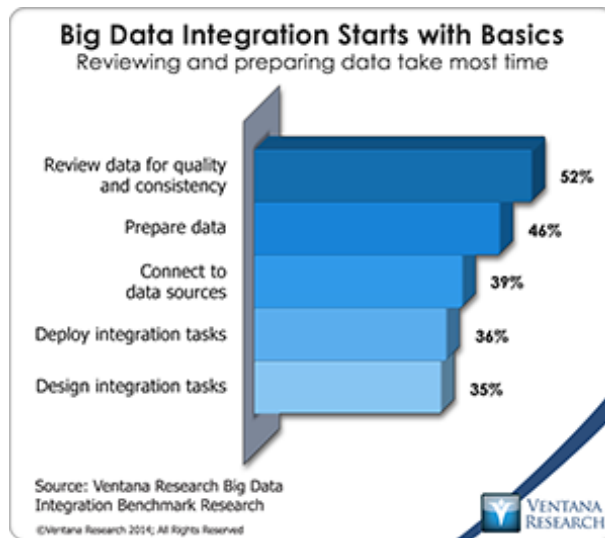©Ventana Research 2014; All Rights Reserved

Establishing a big data environment requires integrating data through proper preparation and potentially continuous updates of data, whether in real time or batch processing. A further complication is that many organizations will not have only one but several big data environments to be integrated into the overall enterprise architecture; that requires data and systems integration. Our research finds that some organizations are aware of this issue: Automating big data integration is very important to 45 percent and important to more than one-third. Automation can not only bring efficiency to big data but also remove many risks of errors or inaccurate and inconsistent data.

Data integration technologies have evolved over the past decade, but advances to support big data are more recent. Our research shows a disparity in how well organizations handle big data integration tasks. Those that are mostly or completely adequate are accessing (for 63%), loading (60%), extracting (59%), archiving (55%) and copying (52%) data while the areas most in need of improvement are virtualizing (39%), profiling (37%), blending (34%), master data management (33%) and masking for privacy (33%). At the system level, the research finds that 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%). 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%) said they will use an internal sandbox with specialized big data. This group of findings reveals that big data integration has enterprise-level requirements that go beyond just loading data to build on advances in data integration.

Big data must not be a separate store of data but part of the overall enterprise and data architecture; that is necessary to ensure full integration and use of the data. Organizations that see data integration as critical to big data are embarking on sophisticated efforts to achieve it. The data integration capabilities most critical to 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%).

Big Data Integration Starts with Basics
Reviewing and preparing data take most time

Review data for quality and consistency — 52%
Prepare data — 46%
Connect to data sources — 39%
Deploy integration tasks — 36%
Design integration tasks — 35%

Source: Ventana Research Big Data Integration Benchmark Research
©Ventana Research 2014; All Rights Reserved

VENTANA RESEARCH

Other organizations are still examining how to automate integration tasks. The most common barriers to improving big data integration are cost of the software or license (for 44%), lack of resources to use on improvement (37%) and the sense that big data technologies are too complicated to integrate (35%). These findings demonstrate that many organizations need to better understand the efficiency and cost savings that can be realized by using purpose-built technology instead of manual approaches using tools not designed for big data. 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. The most time spent in big data integration today is on basic tasks: reviewing data for quality and consistency (52%), preparing data for integration (46%) and connecting to data sources for integration (39%). The first two are related to ensuring that data is ready to load into big data environments. Data preparation is a key part of big data and overall information optimization. More vendors are developing dedicated technology to help with it.

For a process as complex as big data integration, choosing the right technology tool can be difficult. More than half (55%) of organizations are planning to change the way they assess and select such technology. Evaluations of big data integration tools should include considerations 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 have no preference. Half of organizations plan to use cloud computing for managing big data; another one-third (32%) don't know whether they will. The research shows that the most important technology and vendor criteria used to evaluate big data integration technology are usability (very important for 53%), reliability (52%) and functionality (49%). These top three evaluation criteria are followed by manageability, TCO/ROI, adaptability and validation of vendors. Organizations are most concerned to have technology that is easy to use and can scale to meet their needs.

Big data cannot be used effectively without integration; we observe that the big data industry has not paid as much attention to information management as it should – after all, this is what enables automating the flow of data. Organizations trying to use big data without a focus on information management will have difficulty in optimizing the use of their data assets for business needs. Our research into big data integration finds that the proper technology is critical to meet these needs. We also learned from our benchmark research into big data analytics that data preparation is the largest and most time-consuming set of tasks that needs to be streamlined for best use of the analytics that reveal actionable insights. Organizations that are initiating or expanding their big data deployments whether onpremises or within cloud computing environments should have integration at the top of their priority list to ensure they do not create silos of data that they can't fully exploit.

Mark Smith is CEO and chief research officer at Ventana Research.

JOIN THE DISCUSSION

Comment

SEE MORE IN

Big Data/Analytics

RELATED TAGS

[Big Data](#),
[Analytics](#)
Comments (0)

Be the first to comment on this post using the section below.

Add Your Comments:

```
Add your comments here.
```

☐ Notify me when other readers comment on this article.

[Click here to receive notifications without commenting](#)

POST ▶

CANCEL

Videos

[60 Seconds Smarter: E&M Guidelines Explained](#)

Most Read
Most Emailed

- [10 Internet of Things (IoT) Trends & Considerations](#)
- [Data Acceleration: Turning Technology Into Solutions](#)
- [10 Tech Trends Through 2015](#)

- [Next Shift: From Big Data to Deep Data](#)
- [Data Science And Statistics: Colleges Must Evolve](#)

Featured Topics

- [BIG DATA](#)
- [BUSINESS INTELLIGENCE](#)
- [DATA INTEGRATION](#)
- [DATA MGMT](#)
- [DATA MODELING](#)
- [DATA QUALITY](#)
- [FINANCIAL SERVICES](#)
- [GOVERNANCE, RISK & COMPLIANCE](#)
- [HEALTH CARE](#)

[Slideshows](#)

[10 Big Data Career Killers](#)

[Gartner's Top 10 Cloud Myths](#)

[Gartner's Top 10 Strategic Tech Trends for 2015](#)

[8 Principles for Adopting Casework for Data Governance](#)

[Debunking 5 of the Biggest Data Myths](#)

[Five Habits of Data-Driven People](#)

[More Slide Shows »](#)

Current Buzz on Analytics

Twitter

Facebook

LinkedIn

- about 2 days ago we were checking out Get Ready for BI Change bit.ly/14MbHk3
- about 2 days ago we were checking out Cloud vs. Enterprise CRM Software Price War Coming bit.ly/1xzw0Ik
- about 2 days ago we were checking out Obama Touts Well-Wired Iowa Town as Model for City-Run Web bit.ly/14LVone
- about 3 days ago we were checking out Obama Cybersecurity Plan Could Change Healthcare Processes bit.ly/159Q0uL
- about 3 days ago we were checking out Enterprise Marketing's Mobile Mind Shift bit.ly/14FVxsd
- about 4 days ago we were checking out Cloud Security Concerns Shift From IT to Boardroom bit.ly/1FWbSun
- about 5 days ago we were checking out Large Enterprises Embrace Disaster Recovery as a Service (DRaaS) bit.ly/1ya8LtE
- about 6 days ago we were checking out Data Acceleration: Turning Technology Into Solutions bit.ly/14SDAXX
- about 6 days ago we were checking out Next User Experience Battleground: Mobile Customer Experience bit.ly/1FNn4t4
- about 9 days ago we were checking out CNA Using Big Data to Fight Insurance Fraud bit.ly/1y0PSI3
- about 10 days ago we were checking out Five Social Data Marketing Resolutions bit.ly/1FvHZ3M
- about 10 days ago we were checking out Personal Healthcare Data: Patients Will Take Control bit.ly/1AMBkhj

- [about 11 days ago](#) we were checking out Next Shift: From Big Data to Deep Data [bit.ly/1xT9YUz](#)
- [about 11 days ago](#) we were checking out Global Tech Market Looking Better For 2015 [bit.ly/143GKY1](#)
- [about 11 days ago](#) we were checking out Data Science And Statistics: Colleges Must Evolve [bit.ly/1FpdFrs](#)
- [about 11 days ago](#) we were checking out Flash Storage In the Data Center Catching On [bit.ly/14qb1B4](#)
- [about 11 days ago](#) we were checking out CNA Using Big Data to Fight Insurance Fraud [bit.ly/14pO6Wl](#)
- [about 13 days ago](#) we were checking out Patient Monitoring Market: $5 Billion by 2020 [bit.ly/1FgARbm](#)
- [about 13 days ago](#) we were checking out Why Once-Successful Companies Fail [bit.ly/1vYZoHn](#)
- [about 13 days ago](#) we were checking out Average Mobile App Earns "D" Grade From Customers [bit.ly/1vTsrMv](#)

[Follow @infomgmt](#) ⟨ 9,942 followers

[Analytics](#)

- [Hail to the Chiefs! Meet the CDO, CAO and CSO](#)
- [IBM: New Mainframe Boasts Improved Ability to Perform Analytics](#)
- [Data Acceleration: Turning Technology Into Solutions](#)
- [CNA Using Big Data to Fight Insurance Fraud](#)

[Business Intelligence](#)

- [Get Ready for BI Change](#)
- [UPS ORION Advanced Analytics Case Study](#)
- [Business Intelligence Analysts as Architects](#)
- [The Three C's of Self-Service BI](#)

[Customer Experience](#)

- [Insurance Executives See IT Role In Customer Satisfaction](#)
- [Cloud Security Concerns Shift From IT to Boardroom](#)

- [Next User Experience Battleground: Mobile Customer Experience](#)
- [Average Mobile App Earns "D" Grade From Customers](#)

Open Source

- [CNA Using Big Data to Fight Insurance Fraud](#)
- [Hadoop Isn't for Everyone](#)
- [China Plans Foreign Technology Purge, Shift to Homegrown Suppliers](#)
- [8 Hadoop Big Data Predictions for 2015](#)

Predictive Analytics

- [Know More, Sell More With Data Science in 2015](#)
- [Internet of People, Not Things, Is Next Frontier](#)
- [The MIT Analytics Imperative](#)
- [Google's Big Data to Help Auction.com Predict Homebuying Trends](#)

Data Governance

- [CIO Priorities: Strategic Architecture vs Quick Project Wins](#)
- [Obamacare Dental Heathcare Data Doesn't Add Up](#)
- [Target Wants Data Breach Bank Claims Dismissed](#)
- [Creating the Data Governance Killer App](#)

Data Integration

- [Big Data Requires Integration Technology](#)
- [Actuate, Oracle, Red Hat, Others Announce New or Updated Products](#)
- [How Healthcare CIOs View the Data Analytics Landscape](#)
- [Headache Number Two: Remedies for Data Migration Pain](#)

Data Management

- [Large Enterprises Embrace Disaster Recovery as a Service (DRaaS)](#)
- [Personal Healthcare Data: Patients Will Take Control](#)
- [Why Once-Successful Companies Fail](#)
- [Top Healthcare Technology Stories of 2014](#)

- [HOME](#)
- [About Us](#)
- [Contact Us](#)
- [Content Licensing](#)
- [Advertise with Us](#)
- [Customer Service](#)
- [Feedback](#)
- [My Account](#)
- [Site Map](#)
- [Privacy Policy](#)
- [Editorial Submissions](#)

[**sourcemedia
corporate site**](#)

banking

- [American Banker](#)
- [Bank Technology News](#)
- [American Banker Magazine](#)
- [Credit Union Journal](#)

MORTGAGES

- [National Mortgage News](#)

PAYMENTS

- [PaymentsSource](#)
- [Collections & Credit Risk](#)
- [ISO & Agent](#)

capital markets

- [Mergers & Acquisitions](#)
- [Asset Securitization Report](#)
- [Leveraged Finance News](#)
- [Private Placement Letter](#)
- [Traders Magazine](#)
- [Securities Technology Monitor](#)

MUNICIPAL FINANCE

- [The Bond Buyer](#)

accounting

- [Accounting Today](#)
- [Tax Pro Today](#)

HEALTHCARE & BENEFITS

- [Employee Benefit News](#)
- [Employee Benefit Adviser](#)
- [Health Data Management](#)
- [Insurance Networking News](#)
- [Information Management](#)

INVESTMENT ADVISORY

- [Financial Planning](#)
- [On Wall Street](#)
- [Bank Investment Consultant](#)
- [Money Management Executive](#)
- [Investment Mgt. Mandate Pipeline](#)