Big Data Executive Survey

Creating a Big Data Environment to Accelerate Business Value

NVP
NewVantage Partners

Foreward | Thomas H. Davenport

I'm pleased that NewVantage Partners conducted this survey and made the results available to me for review prior to publication. There is a lot of interest in the topic of "Big Data" as evidenced by the high response rate to the survey. The survey yielded some interesting findings, many of which are relevant to research I have conducted. Here are several that are worthy of comment:

There is more going on in large firms in the Big Data space than I would have expected. Both the press and my research suggest a high level of Big Data activity among startups, but less among large companies. This may mean that financial services companies, which were heavily represented in the survey, are ahead of most other firms in their exploitation of Big Data. It may also mean that the respondents didn't make a strong distinction between Big Data and traditional analytics projects.

The results suggest—and I already believed—that Big Data is not a terribly useful term. For the survey respondents, the variety of data sources is a more salient attribute than the volume of the data. Perhaps we should call it "mashup data" instead.

My research on Big Data startups suggests that the primary objective is less to support internal decision-making, and more to support new customer-facing products and processes. There is some support for this in the survey; three years from now, the most common benefit expected is an improved customer experience. Now, however, it's better decision-making.

Regardless of whether the data are supporting decisions or products/processes, the primary beneficiary of Big Data appears to be the customer. The single most common application of Big Data in this survey is probably—reading between the lines of several question responses—the use of multiple customer data sources to better understand their needs and target promotions and offers to them.

Big Data is often described as unstructured, but these responses suggest that structured transactions are the most common data source. In many cases these are probably being combined with less structured data like web clickstreams and social media content. Since the respondents are focused on integrating various sources and types of data, **data integration will continue to be one of the greatest challenges faced by IT organizations.**

There is a storm approaching on the Big Data talent front; 70% say they plan to hire data scientists, but they already find this "challenging" to "extremely difficult," and there is no reliable source of new talent in this category. It would seem to be a wise move to begin "building" such talent as well as "buying" it. In addition, the survey results suggest that organizations need to provide training to executives and decision- makers on how Big Data is going to change their businesses, and how to manage it.

This sampling of Big Data activity is one of the few I have seen that focuses on large organizations and offers responses from C-level executives. It should be very useful to organizations in suggesting immediate and longer-term actions for managing this valuable new resource.

Thomas H. Davenport is a Visiting Professor at Harvard Business School and the co-founder and Director of Research at the International Institute for Analytics. He is also the author of the highly acclaimed Competing on Analytics: The New Science of Winning, cited by CIO Magazine as one of the all-time "Top 15 Most Groundbreaking Management Books."

Introduction

In the Spring of 2012 at a NewVantage Partners Executive Thought Leadership Event, a group of leading business executives asked us to explore the issues and business trends in America's leading companies relating to "Big Data." During the summer of 2012, NewVantage Partners conducted an exclusive survey with C-level executives and function heads from many of America's leading companies. This in-depth industry survey targeted the most senior Fortune 500 and Federal Agency business and technology leaders responsible for overseeing enterprise Big Data initiatives.

The survey identified five core themes and trends, which were the highlight of our first survey report issued in September 2012. These themes and trends were also core to follow-up discussions we had with over 40 of the companies that responded to the survey and an additional dozen firms that subsequently shared their Big Data efforts with us. They were:

- Promise of Improved Data-Driven Decision Making
- Customers and Risk are the Primary Focus
- It's About Variety, Not Volume
- Organizational Alignment is Critical to Success
- Biggest Roadblock May Be More about People Than Technology

This final trend was the topic of our second survey report published in October 2012, ***People – The Key to Success***, which addressed the broad reaching talent challenges facing Big Data in the enterprise.

This is our final report and focuses on what we learned from both the data gathered in the survey and our subsequent briefings with respondents. It combines an analysis of the survey findings with a synthesis of the state of Big Data in the enterprise and how organizations should planning for successful implementation of their initiatives.

The potential power of Big Data technology is clear, but the path to success for companies is neither simple nor straightforward. It is complex and requires careful planning and clear roadmaps to offer the highest likelihood of success. The key is to "Create a Big Data Environment to Accelerate Business Value."

We are pleased to share with you this final report in a series of observations from the survey. To receive a copy of ***Part I: Themes and Trends or Part II: People – The Key to Success***, simply send a request to bigdata@newvantage.com.

Methodology

The survey posed 65 in-depth questions designed to provide a useful benchmark for enterprises seeking to understand the state of Big Data initiatives among peer institutions, and to answer critical questions such as:

- How much are enterprises investing in Big Data initiatives?
- Where is the sponsorship and funding for Big Data initiatives coming from?
- What are the initial applications, kinds of data, and solutions/approaches that enterprises are employing for their Big Data initiatives?
- Why is Big Data an important initiative for leading enterprises?
- Where do organizations stand in terms of the comparative maturity of their Big Data initiatives and their rate of progress?
- Which kinds of enterprises are at the forefront of Big Data capabilities and initiatives?

Please note that in certain cases, the total percentage in a table does not always add up to exactly 100%, but a number very close to it (e.g., 99 or 101), because of rounding. When questions allow for several responses, percentages often add up to more that 100%.

Demographics

Survey respondents included Chief Data Officers, Chief Information Officers, Chief Technology Officers, Chief Analytics Officers, Chief Information Architects, Line-of-Business Heads, and senior Function Heads (SVP/VP). The breakdown of respondents was as follows:

- C-Suite Executives – 30%
- Enterprise Function Heads (EVP/SVP) – Head of Analytics, Head of Informatics, Head of Enterprise Information – 50%
- Data Program Head (SVP/VP) – Head of Data Architecture, Chief Data Architect – 20%

Organizationally and functionally, the roles of the respondents were:

- IT – 47%
- Business Analytics – 30%
- Line-of-Business Management – 23%

Over fifty executives representing leading Fortune 1000 companies and large Federal agencies participated in the survey. All participants were executives with budgetary and decision-making responsibility or direct visibility and influence for Big Data initiatives.

Financial service companies have traditionally been at the forefront of using strategic data and analytics to support their core business functions and they were the most strongly represented industry group. Industry representation is summarized as follows:

- Financial Services – 53%
- Insurance – 19%
- Government – 10%
- Other Business – 18%

The majority of these organizations are very large, with at least 30,000 employees, and they operate diverse businesses in multiple markets. A complete list of participating institutions is provided in Appendix A.

We wish to thank all those individuals and organizations that participated in the survey. Individual respondents and individual company responses will remain anonymous.

Creating a Big Data Environment to Accelerate Business Value

"Big data is out there. The trick is finding ways to make it work for your company."

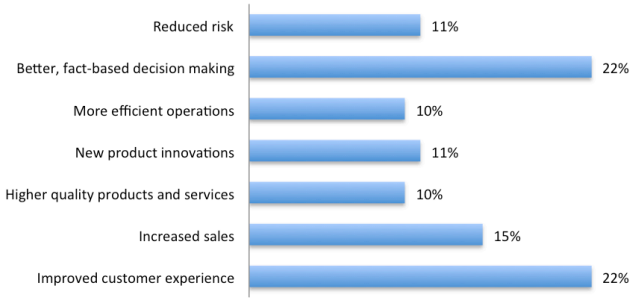
– Adi Ignatius, Editor-in-Chief
Harvard Business Review

Improve Data-Drive Decision Making

We are living in the era of "Big Data." This is a world of vast opportunities, where huge volumes of disparate data from a broad range of new and existing sources can be mined to deliver business decisions with unprecedented speed, automation, and intelligence. This is a game changer for everything from marketing programs to risk and fraud applications.

When asked to rank the tangible benefits they hope to achieve through Big Data, survey respondents offered a broad range of responses, but there were two clear leaders: ***Better, fact-based decision making*** and ***Improved customer experience***. [Figure 1] The theme of "The Promise of Improved Data-Driven Decision Making" was clearly identified as a major focus of the organizations that participated in the survey. 85% of the participating companies have Big Data initiatives underway and the primary reason for these investments is to improve analytic capabilities and make smarter business decisions. In our subsequent briefings with respondents, however, we learned that in addition to making better decisions, the real quantum leap for companies is accelerating the speed at which they can get to a decision. This something we refer to as "Time-to-Answer" (TTA).

Figure 1
What tangible benefits do you hope to achieve through your Big Data initiatives? (rank all that apply)

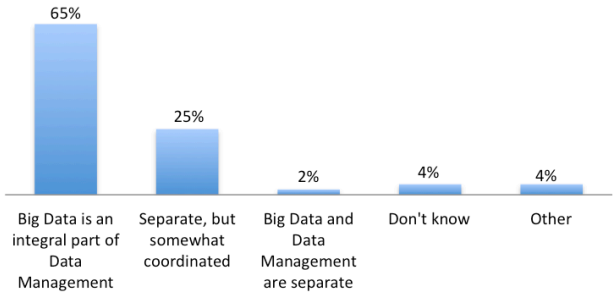


Most of the executives who participated in the NewVantage Partners Big Data Survey see these opportunities in context. They understand that leveraging Big Data to deliver real business value will require a carefully focused strategy that leverages and protects existing data assets, develops new capabilities that are production-ready and reusable, and is capable of managing the deluge of new data that is created every day. To succeed, methodical and careful planning is a must.

Big Data, Analytics, and Organizational Alignment

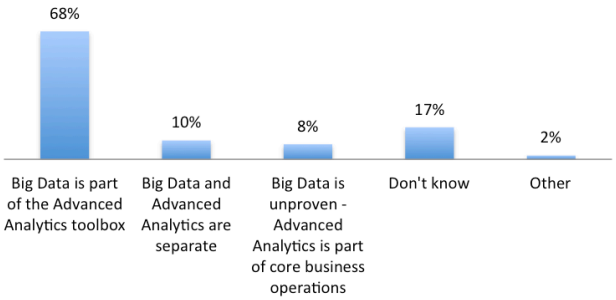
Historically, there has been much talk about the difference between traditional analytics and Big Data, and organizational responsibility for each within an enterprise. The survey, however, shows the two are becoming closely intertwined and must work together to deliver the promised results of Big Data. Further, breaking down organizational boundaries and creating close integration between IT organizations and the business units is a critical step for any organization hoping to build a winning strategy for Big Data. Data management and analytics have often resided in different parts of the organization. IT departments usually controlled the data and analytics was conducted in either a special group or within a business unit. This is contrary to the entire principal of Big Data and the survey confirms that organizations understand close integration is necessary. 65% felt "Big Data is an integral part of Data Management," and 68% further felt that "Big Data is part of the Advanced Analytics toolbox." [Figure 2 & Figure 3]

Figure 2
How are you thinking about Big Data with respect to overall Data Management? (check one)



NVP
NewVantage Partners

Figure 3
How are you thinking about Big Data capabilities with respect to Advanced Analytics (data mining, predictive modeling, etc.) initiative? (check one)



This shift in mindset is a critical step in creating organizational initiatives that will allow Big Data to make a meaningful business impact. Integrating real-time, full analytic capabilities into the business and operating units will enable the type of quick reactions to key business questions and challenges that can build competitive advantage and improve performance. As pointed out by Thomas Davenport, Paul Barth and Randy Bean in their recent MIT Sloan Management Review article, "How Big Data is Different," change is a must:

> *"Coming to terms with big data is prompting organization to rethink their basic assumptions about the relationship between business and IT – and their respective roles . . . Whereas the most vaunted business and IT capabilities used to be scalability and scale, the new advantages are based on discovery and agility – the ability to mine existing and new data sources continuously for patterns, events and opportunities . . . This requires a sea change in IT activity with in organizations . . . As big data evolves, the architecture will develop into an information ecosystem: a network of internal and external services continuously sharing information, optimizing decisions, communicating results and generating new insights for business."*

Accelerating Time-To-Answer: The Real Value of Big Data

One of the clearest messages from the survey was the desire of companies to make better business decisions, faster. We refer to this critical measure of analytical prowess as Time-to-Answer (TTA). Organizations are using Big Data platforms to answer questions in seconds rather than days, and in days rather than months. Accelerating TTA can enable organizations to answer questions that have stubbornly resisted analysis, develop test and learn processes that quickly adapt to the market, and automate complex workflows. However, our work has discovered that these leaps in TTA are lost unless a careful process is followed and the relationship between Big Data and traditional solutions are clearly defined and governed.

The core value of data analysis is to provide answers. The accuracy and speed of those answers are critical to their value. It is critical to focus on the question, think of "Big Question, rather than Big Data." Before a company can hope to leverage a strategy to improve TTA, it is important to take a step back and think carefully about the steps involved in preparing to answer a complex business question:

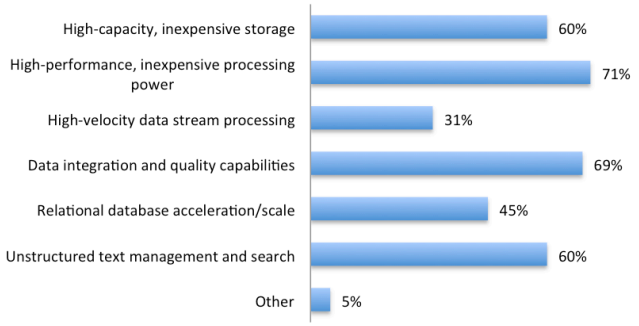
1. Clarify the question and the type of answer needed = build the business case
2. Identify the data required and analysis approach
3. Source the data
4. Cleanse, normalize, and integrate the data
5. Analyze the data
6. Validate the results
7. Present or apply the answer = measure the results

These steps are far from a simple, sequential process. Most companies spend a far larger amount of time (80% or more) on steps 3 and 4 - finding the data and preparing it for analysis - than actually analyzing the data. These steps are critical to accelerating TTA and Big Data solutions enable the process in a few novel ways.

First, the low price and high capacity of big data platforms allow businesses to load *all* that data from source systems, regardless of whether or not it may be needed for the question at hand. While this may seem wasteful, it eliminates two important delays: writing programs to select just the data needed, and going back to the source systems multiple times as new insights generate new questions that need new data. Building traditional data marts and data warehouses is extraordinarily complex and costly. The broad range of open source offerings coupled with flexible, scalable grid systems create an environment that not only drives down costs, but also offers the potential of decreasing query times exponentially. When asked about the data management features and functions that were most important, survey respondents highlighted approaches that offered less

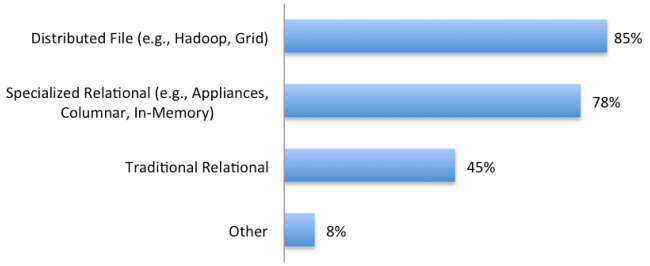
expensive storage and processing, and an ability to improve the quality of data integration and deal with unstructured data. [Figure 4]

Figure 4
What data management functions or features are most important to you? (check all that apply)



Second, big data platforms - unlike relational databases - allow analysts to organize, clean, and integrate data selectively, ignoring records and fields that are not the current focus of analysis. In traditional systems such as data warehouses, significant effort is spent on *data engineering* to ensure that all the data is production-worthy before releasing it to users. Big data techniques, such as key-value modeling, allow analysts to create a data set *ready to use*; full data engineering creates data *ready to re-use*. By deferring full data engineering, big data platforms accelerate TTA during discovery-oriented analysis, and eliminate the engineering effort on data that does not deliver value. A movement away from traditional relational databases is clear. This is by no means abandonment, but the cost of scaling these systems makes Big Data cost prohibitive and doesn't create an environment that can effectively leverage large volumes of disparate data. The focus on Hadoop and Grid is strong, followed closely by appliance and other specialized systems. [Figure 5]

Figure 5
What data management approaches are you considering? (check all that apply)



These emerging technologies and methodologies can have a dramatic impact on accelerating analysis and more quickly providing answers to complex business questions, but many products are new and unproven. Organizations must carefully enter this new environment and make sure they possess the requisite skills and are organized to successfully leverage these new platforms to achieve their goals.

A key to sustainable success in the business world is to learn what you do not know. Successful companies constantly challenge their assumptions and learnings of the past by constantly looking for new insights into their customers, markets, products, and risks. This approach—pursuing the "new" while operating on the "known"—is a healthy continuous improvement model. However, it has an important impact on how companies leverage their data and analytics—and most do it wrong.

Before new data discoveries, organizations must first govern their trusted sources of data. These are the architecture guidelines that should be considered for governing 'known' information:

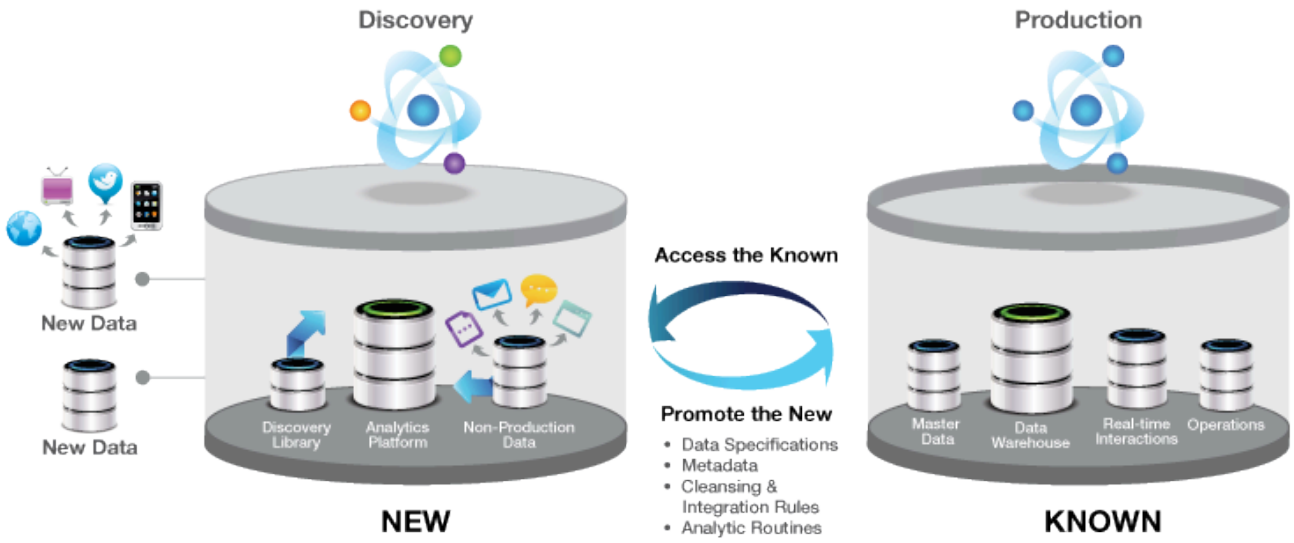
- Cleansed, mastered, and pre-integrated analytic and operational data
- Common metrics, models, and BI capabilities
- Up-to-date business and technical metadata
- Rigorous change management
- *TTA Accelerators*: Standards, scale & high availability

Conversely, the new data generated today can be a bit more unruly. Enabling new corporate insights requires organizations to draw on data from sources both internal and external, including social media, government data, service and sales data, and research. Therefore, "new" discovery must be lightly governed, and include

- Access to known data & definitions
- Secure, large workspace for working with external data
- Rapid profiling, quality, and data integration tools
- Non-standard analytic routines and data structures
- Ability to load complete data sets without data modeling
- *TTA Accelerators*: Powerful infrastructure, specialized tools

NVP NewVantage Partners

Creating a Big Data Environment to Accelerate Time-to-Answer



Enable the New

- Access to known data & definitions
- Secure workspace for working with external data
- Rapid profiling, quality, and data integration tools
- Non-standard analytic routines and data structures
- TTA Accelerators: Powerful infrastructure, specialized tools

Govern the Known

- Cleansed, mastered, and pre-integrated analytic and operational data
- Common metrics, models, and BI capabilities
- Up-to-date business and technical metadata
- Rigorous change management
- TTA Accelerators: Standards, scale & high availability

The "new" and the "known" are not islands; they must be symbiotic systems connected to and feeding each other. "New" analyses require rapid access to all the "known" data representing the reality of today's business. Conversely, there must be a disciplined approach to promoting new insights, data, and models to evolve the "known." Without this linkage, the systems diverge into incoherence that does not reconcile or scale.

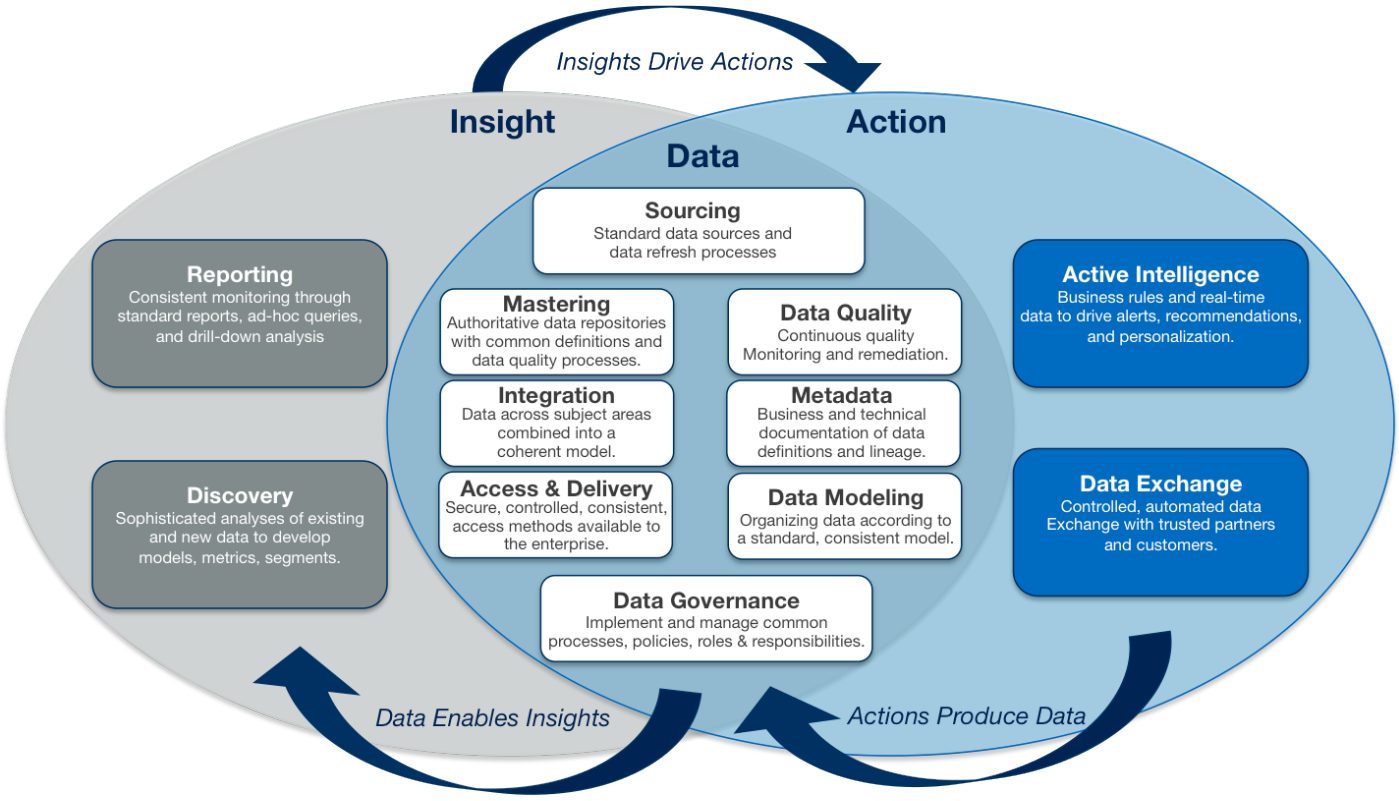
Emerging technologies and methodologies – including Hadoop, Cloudera, database appliances, accelerators and self-learning and genetic algorithms – can dramatically reduce TTA. The key to streamlining the time from a corporate question to game-changing business insight is to right-size the approach to analysis: rapid iterations during discovery, and rigorous engineering into production. Strong governance and the oversight of known data capabilities must coexist with agile data analysis that paves the way for new data discoveries. Enterprises with the capability to create an ebb and flow between dynamic discovery and scalable execution position themselves for sustainable success and dominance over their competitors.

Turning Insights Into Action – Creating An Ecosystem

We are living in a world exploding with Big Data opportunities and one in which the landscape is constantly changing. Big Data is currently at the top of the technology hype cycle at the moment. Whether it is yet another new technology to help companies leverage "Big Data" or the latest article hyping the promise of Big Data to "change the world," to succeed in "Creating a Big Data Environment to Accelerate Business Value," organizations must be rigorous in their preparation and execution of initiatives.

NVP NewVantage Partners

The promise of Big Data is enormous, but it must be taken in context. These new technologies can help dramatically reduce TTA, but enterprises have been managing data and performing analytics for years. At NewVantage Partners, we believe Big Data must be viewed as an ecosystem: capture and create data, clean and organize it, mine business insights from it, and use those insights to drive intelligent business actions. This must be a closed loop that feeds data back into the system to allow companies to use data to test, learn and improves processes.



To provide meaningful and actionable insights, the data must be “good data.” This means there must be established practices for creating and managing high quality enterprise information assets. These include all standard data-management capabilities, such as data sourcing and integration; quality and metadata management; data modeling; and data governance. Insight capabilities include tools, data, and processes for management reporting and advanced analytics. Action capabilities provision data and business intelligence to applications, business processes, and business partners, and capture responses to interactions. Big Data enables companies to more effectively deal with a larger volume and wider variety of data, but it does not address or solves issues around the quality of that data. Organizations must continue to focus on this important data management effort.

Big Data presents both opportunities and challenges in each of these domains. Data management leverages Big Data technology to eliminate redundancy and provide scalable infrastructure for managing Big Data assets. Insight uses appliances and accelerators, NOSQL technology, and automated analytics to expose new value hidden in big data. Businesses deploy these insights through intelligent agents mediating both internal and external communications and interactions.

NVP
NewVantage Partners

Conclusion

Big Data presents organizations with new capabilities for driving business value. Our Executive Survey and resulting discussions with firms that are at the forefront or represent the mainstream of corporate Big Data initiatives have highlighted several dominant themes, which we have sought to explicate in the three summary reports we have produced. We would like to thank again all of those firms that have shared their time and provided us with insight into the approaches they are taking.

In synthesizing all that we have seen and been told during the past six months, we have reached these final conclusions:

- An "Ecosystem" must be created:
  - Know what question you are asking and why.
  - Focus on the quality and integrity of the data. Bad data is always bad data.
  - Turn "insights" into "action" and be prepared to measure the results and refine or modify actions as necessary.

- Don't focus too much on the "Big" of Big Data. It's really not about volume. The key is to make better decisions, faster. Sometimes this may mean working with less data, but the right data.

- When sifting through the "hype" around Big Data, remember there is a distinction between what's going on in the enterprise today and the experience of, and initiatives underway at, companies that were born into Big Data, like Facebook, Google, LinkedIn, eBay, Yahoo, etc. These companies have powerful Big Data efforts underway, but they have never faced the challenge of dealing with older legacy data and systems.

- Big Data offers huge potential for business gains. New technologies are more powerful, yet less costly and there is a real chance to improve decision-making and act more quickly, but it is critical to remember that the path to success will not be simple. It will require careful planning, clear definition of goals, hard work, and a willingness to embrace organizational change.

- The traditional three V's associated with Big Data – Volume, Variety, and Velocity – must be expanded for some industries that require a higher degree of accuracy with results. Several companies we have spoken with include Veracity or Validity to carry on the "V" theme. Regulatory, risk and compliance projects have a miniscule margin of error, as do life science and healthcare projects, while a consumer marketing campaign can tolerate a higher margin of error.

Appendices

Appendix A – Participating Companies

Banking

Bank of America
Bank of New York Mellon
CitiGroup
JP Morgan
RBS Citizens Financial
State Street Bank
US Bank
Wells Fargo Bank

Insurance | Health Care

Aetna
Broad Institute
Cigna
CVS/Caremark
The Hartford
Harvard Pilgrim Health Care
SunLife Financial
Travelers
United Healthcare

Investments

Charles Schwab
Conning Asset Management
Fidelity Investments
ING
TD Ameritrade
TIAA-CRFF
Wellington Financial

Other Business and Financial Services

American Express
Freddie Mac
General Electric (GE)
MasterCard
Pitney Bowes
Thomson Reuters
VISA
Wright Express

Media and Technology

Avid Technology
Time Warner Cable

Government

Department of Defense (DOD)
General Services Administration (GSA)
Department of Health and Human Services (HHS)
Social Security Administration (SSA)