Research Report Abstract

Enterprise Big Data, Business Intelligence, and Analytics Trends: *Redux*

By Nik Rouda, ESG Senior Analyst; and Bill Lundell, Director of Research; With Jennifer Gahm, Senior Project Manager

July 2016
Introduction

Research Objectives

There is still confusion in the market about the current state of data initiatives, especially when it comes to the potential impact of big data on traditional business intelligence and analytics practices. Customers and vendors alike are working to build robust solutions even as new approaches, technologies, and best practices continue to evolve rapidly. Still, few can accurately assess the motivations, impacts, and implications of the shifts. ESG undertook a broad study to establish a baseline on various data initiatives and their relative maturity.

While many big data vendors offer “one size fits all” platforms, the benefits of big data can be magnified in larger businesses, where it can affect more workers and have impact on a bigger scale. This portion of the addressable market may face unique needs, concerns, and opportunities. What is important is to understand how these groups are different from smaller businesses. Understanding these differences will help vendors to be more successful in providing appropriate offerings.

In order to assess the overall status of data initiatives in public and private organizations, ESG surveyed 200 IT and business professionals representing enterprise (1,000 or more employees) organizations in North America. All respondents were personally responsible for or are familiar with their organizations’ current database, business intelligence (BI), and/or analytics solutions.

The survey was designed to answer the following questions:

- Relative to all other business and IT priorities over the next 12-18 months, how do organizations rate the importance of their big data analytics projects and initiatives?
- For new initiatives in the area of big data and analytics, how long do respondents think it will take for their organization to start seeing significant business value?
- How much data is processed as part of a typical data analytics exercise?
- What data analytics challenges are organizations facing?
- What is the sentiment for Spark, and what is driving interest in the technology?
- What level of interest exists for Hadoop? What is the anticipated impact of Hadoop on the traditional data warehouse approach?
- What type of Hadoop distribution(s) are organizations currently evaluating? How important is open source or Open Data Platform support/participation to an organization’s choice of Hadoop distribution(s)?
- What will likely be the primary deployment strategy for net-new big data and analytics deployments going forward?
- What are the perceived advantages associated with the usage of cloud-based big data and analytics solutions?

Survey participants represented a wide range of industries including manufacturing, financial services, health care, communications and media, retail, government, and business services. For more details, please see the Research Methodology and Respondent Demographics sections of this report.
## Contents

- List of Figures ........................................................................................................................................................................3
- List of Tables ...........................................................................................................................................................................3
- Executive Summary .................................................................................................................................................................4
- Report Conclusions .............................................................................................................................................................4
- Introduction ............................................................................................................................................................................5
- Research Objectives ............................................................................................................................................................5
- Research Findings ...................................................................................................................................................................6
- Enterprise organizations are making tangible progress on their big data initiatives .................................................................6
- Enterprise big data volumes are starting to live up to their name, but integration is still the biggest challenge ..................8
- Spark already rivals Hadoop, though Hadoop is making headway in the data warehouse space .......................................11
- “Openness” is highly valued when it comes to Hadoop usage and consideration .................................................................15
- Big data deployments are getting cloudy in the enterprise .............................................................................................17
- Conclusion .............................................................................................................................................................................20
- Research Methodology .........................................................................................................................................................21
- Respondent Demographics ...................................................................................................................................................22
- Respondents by Primary Area of Responsibility ...............................................................................................................22
- Respondents by Age ..........................................................................................................................................................22
- Respondents by Age of Organization ................................................................................................................................23
- Respondents by Number of Employees ............................................................................................................................23
- Respondents by Industry ..................................................................................................................................................24
- Respondents by Annual Revenue .....................................................................................................................................24
List of Figures

Figure 1. Importance of Big Data Analytics Projects and Initiatives Relative to Business and IT Priorities ........................................... 7
Figure 2. Expected Length of Payback Period for New Big Data and Analytics Initiatives ................................................................. 7
Figure 3. Amount of Data Processed During a Typical BI/Analytics Exercise ......................................................................................... 9
Figure 4. Amount of Data Processed During a Typical BI/Analytics Exercise, 2014 versus 2015 ................................................................ 9
Figure 5. Data Analytics Challenges by Average Number of Unique Data Sources .................................................................................. 10
Figure 6. Hadoop Implementation Plans ............................................................................................................................................ 12
Figure 7. Spark Implementation Plans ................................................................................................................................................ 13
Figure 8. Factors Driving Interest in Spark ....................................................................................................................................... 13
Figure 9. Most Important Features When Considering Solutions to Use SQL on Hadoop .............................................................. 14
Figure 10. How Current Users Anticipate Hadoop Affecting Traditional Data Warehouse Approach ............................................................. 14
Figure 11. Type of Hadoop Distribution Planned Adopters Are Evaluating .................................................................................... 15
Figure 12. Importance of Open Source Support/Participation in Choice of Hadoop Distribution ......................................................... 16
Figure 13. Type of Hadoop Distribution Planned Adopters Are Evaluating .................................................................................... 16
Figure 14. Primary Deployment Strategy for Net-new Big Data and Analytics Deployments ............................................................... 18
Figure 15. Preferred Model of Cloud-based Services .......................................................................................................................... 19
Figure 16. Advantages for Cloud-based Big Data and Analytics Solutions ......................................................................................... 19
Figure 17. Respondents’ Primary Area of Technology Responsibility .................................................................................................... 22
Figure 18. Respondents by Age .......................................................................................................................................................... 22
Figure 19. Respondents by Age of Organization ............................................................................................................................... 23
Figure 20. Respondents by Number of Employees ........................................................................................................................... 23
Figure 21. Respondents by Industry .................................................................................................................................................. 24
Figure 22. Respondents by Annual Revenue .................................................................................................................................... 24

List of Tables

Table 1. Data Analytics Challenges by Average Number of Unique Data Sources .............................................................................. 10
Research Methodology

To gather data for this report, ESG conducted a comprehensive online survey of IT and business professionals working at enterprise-class organizations (i.e., 1,000 or more employees) from both private- and public-sector organizations in North America (United States and Canada) between October 22, 2015 and November 2, 2015. To qualify for this survey, respondents were required to be IT or business professionals familiar with their organization’s current big data, database, data warehouse, business intelligence (BI), and/or analytics solutions, as well as forward-looking strategies. All respondents were provided an incentive to complete the survey in the form of cash awards and/or cash equivalents.

After filtering out unqualified respondents, removing duplicate responses, and screening the remaining completed responses (on a number of criteria) for data integrity, we were left with a final total sample of 200 IT and business professionals.

Please see the Respondent Demographics section of this report for more information on these respondents.

Note: Totals in figures and tables throughout this report may not add up to 100% due to rounding.
Respondent Demographics

The data presented in this report is based on a survey of 200 qualified respondents. Figures 1-6 detail the demographics of the respondent base, including individual respondents’ primary area of responsibility and age, as well as respondent organizations’ total number of employees, primary industry, and annual revenue, among others.

Respondents by Primary Area of Responsibility

Respondents’ primary area of responsibility is shown in Figure 1.

Figure 1. Respondents’ Primary Area of Technology Responsibility

Which of the following best describes your primary area of responsibility? (Percent of respondents, N=200)

- Manager of development or developer of business intelligence/analytics solutions, 25%
- Database administrator, 23%
- Data scientist, 5%
- Enterprise or data architect, 6%
- Business analyst, 8%
- Data engineer, 12%
- Data warehouse/business intelligence/analytics manager, 10%
- Data analyst, 12%
- Business analyst, 8%
- Enterprise or data architect, 6%
- Manager of development or developer of business intelligence/analytics solutions, 25%

Source: Enterprise Strategy Group, 2016

Respondents by Age

Figure 2 shows the age of respondents.

Figure 2. Respondents by Age

Please select your age group. (Percent of respondents, N=200)

- 25 and under, 4%
- 26 to 35, 38%
- 36 to 45, 34%
- Over 55, 9%
- 46 to 55, 16%

Source: Enterprise Strategy Group, 2016
Respondents by Age of Organization

Figure 3 shows the age of respondent organizations (i.e., length of time an organization has been in existence).

Figure 3. Respondents by Age of Organization

<table>
<thead>
<tr>
<th>Age of Organization</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 1 year</td>
<td>1%</td>
</tr>
<tr>
<td>1 to 5 years</td>
<td>3%</td>
</tr>
<tr>
<td>6 to 10 years</td>
<td>22%</td>
</tr>
<tr>
<td>11 to 20 years</td>
<td>36%</td>
</tr>
<tr>
<td>21 to 50 years</td>
<td>20%</td>
</tr>
<tr>
<td>More than 50 years</td>
<td>19%</td>
</tr>
</tbody>
</table>

For approximately how long has your current employer been in existence? (Percent of respondents, N=200)

Source: Enterprise Strategy Group, 2016

Respondents by Number of Employees

The number of employees in respondents’ organizations is shown in Figure 4.

Figure 4. Respondents by Number of Employees

<table>
<thead>
<tr>
<th>Number of Employees</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000 to 2,499</td>
<td>44%</td>
</tr>
<tr>
<td>2,500 to 4,999</td>
<td>4%</td>
</tr>
<tr>
<td>5,000 to 9,999</td>
<td>16%</td>
</tr>
<tr>
<td>10,000 to 19,999</td>
<td>4%</td>
</tr>
<tr>
<td>20,000 or more</td>
<td>9%</td>
</tr>
<tr>
<td>1,000 to 2,499</td>
<td>44%</td>
</tr>
</tbody>
</table>

How many total employees does your organization have worldwide? (Percent of respondents, N=200)

Source: Enterprise Strategy Group, 2016
Respondents by Industry

Respondents were asked to identify their organization’s primary industry. In total, ESG received completed, qualified respondents from individuals in 20 distinct vertical industries, plus an “Other” category. Respondents were then grouped into the broader categories shown in Figure 5.

Figure 5. Respondents by Industry

What is your organization’s primary industry? (Percent of respondents, N=200)

Source: Enterprise Strategy Group, 2016

Respondents by Annual Revenue

Respondent organizations’ annual revenue is shown in Figure 6.

Figure 6. Respondents by Annual Revenue

What is your organization’s total annual revenue ($US)? (Percent of respondents, N=200)

Source: Enterprise Strategy Group, 2016