THE DATA WAREHOUSE EVOLVED: A FOUNDATION FOR ANALYTICAL EXCELLENCE

May 2017

Author: Michael Lock
Vice President & Principal Analyst, Analytics & Business Intelligence

Report Highlights

Alongside the rapid advancements in sophisticated data management technology, the data warehouse has (somewhat) quietly evolved to help support the growth in data complexity and disparity. This report explores a Best-in-Class approach to data management and how companies are prioritizing data technologies to drive business growth and efficiency.

ABERDEEN GROUP

2

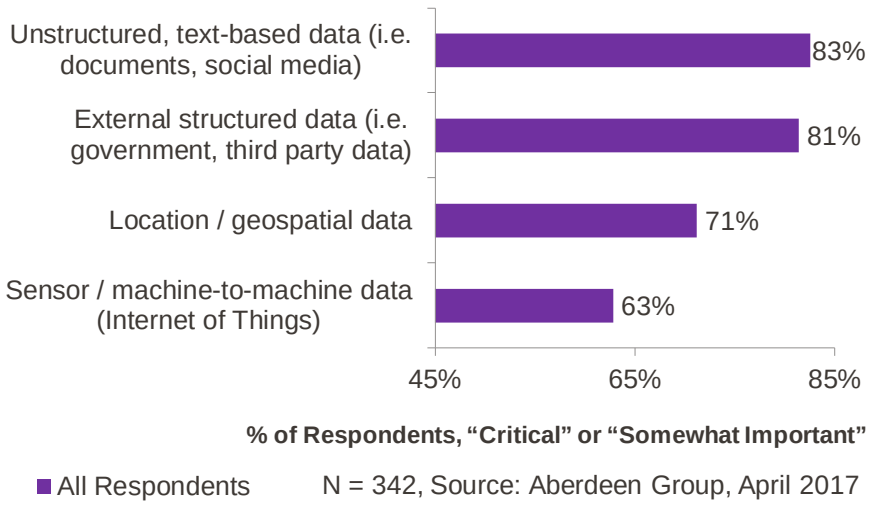
Beyond traditional structured data, companies today are not just storing and managing a wider variety of information, they're exploiting it to power their most important decisions.

Complexity and Disparity — The New Normal

The world of business analytics is evolving rapidly, and while there are multiple emerging trends of note, two stand out as particularly impactful. First, there is an expanding and increasingly diverse audience of users that are becoming more analytically active. From mid-level Line-of-Business staff to senior executives on mahogany row, more users in more job functions are taking an increased level of ownership in the insight that fuels their decisions and the underlying data that supports that insight.

The second trend, closely tied to the expansion and emergence of new users, demonstrates an expanded scope of data in use. Beyond traditional structured data, companies today are not just storing and managing a wider variety of information, they're exploiting it to power their most important decisions (Figure 1).

Figure 1: Breadth of Data Used for Analysis



Most casual observers might consider analytics and business intelligence to be supported exclusively by data sourced from traditional applications like financial data from an ERP system or

Demographics:

This report draws on findings from Aberdeen's **2017 Big Data Survey** of 342 organizations. The pool of respondents was diverse, and broke down as follows:

- Line-of-Business Manager or Staff – 37%
- Executive (VP / C-Level) – 28%
- IT Professional – 22%
- Business Analyst / Data Scientist – 13%

3

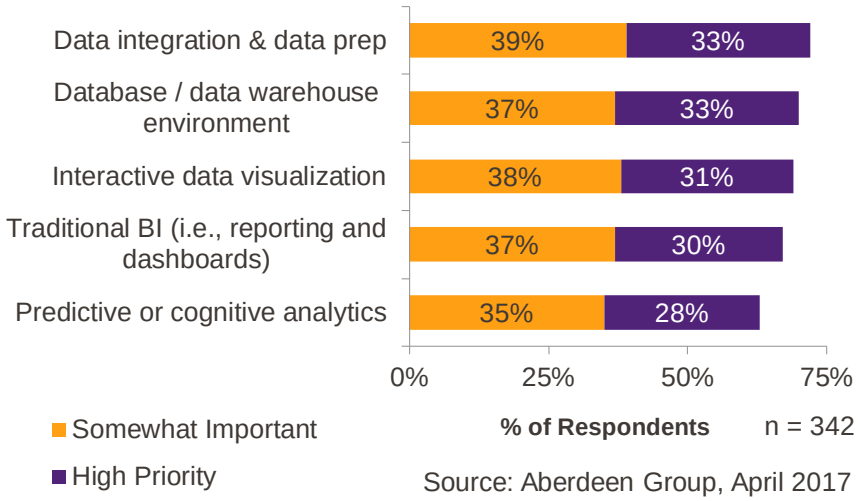
➔ Related Research: *The Hybrid Data Warehouse: Fluid, Flexible, and Formidable*

➔ Related Research: *Data Preparation: The New Normal, the Now Necessary*

customer information from a CRM application. While those types of data likely still support the majority of analytical use cases, companies are increasingly waking up to the potential locked away in other forms of information. Customer sentiment data from social media channels, external third-party data, or machine-generated sensor data all represent pools of potential insight that today's analytically inclined organizations look to tap into. While these data types might not be mission-critical for every company, most companies view all of them as at least somewhat important to their analyses.

Moreover, while there have been significant advancements in front-end visualization and discovery technology, Aberdeen's research demonstrates that foundational data-related technologies currently top the list of investment priorities.

Figure 2: Top Investment Priorities



Source: Aberdeen Group, April 2017

Aberdeen's research demonstrates that organizations are struggling to manage an average of 33 unique data sources used

Despite the hype around open source Hadoop-based data management systems, many companies still gravitate toward the familiarity of a data warehouse, but with the extensibility and flexibility to support today's complex data environments.

for analysis. This type of disparity underscores the need for data integration and preparation technology to normalize the information and make it more consumable for the analytical systems and the people who would use it. Almost equally important for today's companies, however, is an evolved form of the traditional data warehouse technology. Despite the hype around open source, Hadoop-based data management systems, many companies still gravitate toward the familiarity of a data warehouse, but with the extensibility and flexibility to support today's complex data environments.

Profile of a Best-in-Class Company

Any investment in data infrastructure is done with at least a modicum of focus on delivering business results. Companies that leverage data as a strategic asset are able to transform it into insight — and exploit that insight to produce business results. With these ideas in mind, Best-in-Class companies were defined by their ability to perform against the following three key performance metrics, listed below along with the weighted average performance against these metrics:

- **Visibility into key data sources.** Given the demonstrable need for better and more diverse data among the typical user these days, the ability to see and access the right data has become critical.

  62% of Best-in-Class companies report that their visibility into key data is "strong" or "highly effective," compared with 27% of Average companies and only 4% of Laggards.

- **Ability to share data.** With more data visible and accessible to the Line of Business, companies then require the ability to distribute and socialize key information between business functions.

5

➔ Related Research: *MDM without Limits: Connecting Data, Crafting Insight, Mastering Results*

➔ Related Research: *Empowering Analytics: IT and Business Partnerships for Success*

74% of Best-in-Class companies report being "satisfied" or "very satisfied" with their ability to share data across the organization, compared with 43% of Average companies and 14% of Laggards.

➔ **Business efficiency.** An efficient flow of clean and usable data throughout the organization offers a multitude of opportunities to identify and act upon ways of doing business better.

Best-in-Class companies, on average, saw a 20% year-over-year reduction in operating cost, compared with 9% for Average companies, and 6% for Laggard organizations.

The research strongly suggests that the key to elevated performance is marrying technology to the right kind of internal efficiency and maturity to deliver results. Top companies share a variety of capabilities that help them achieve that performance. Internally, the Best-in-Class put a variety of procedures in place that help them empower users with the right data.

First, given the assortment of data sources that could be of use to a decision maker, it doesn't always make sense to force a user to file a formal request with IT for access to a particular application or data source. Best-in-Class companies are more likely to have processes in place, in conjunction with the right kind of user-friendly technology, to accelerate access to new data. Additionally, once that access has been granted and a path has been forged via technology or other process, these top companies recognize the importance of then sharing and socializing that information to other departments and functional areas within the company (Figure 3).

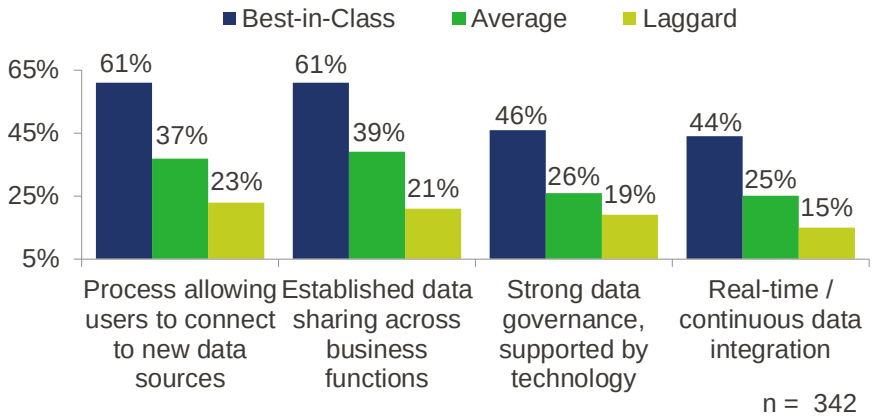
Fast Facts:

Top data-related challenges (% of respondents)

- Difficulty accessing data from different areas of the business (i.e. data silos) – 45%
- Insufficient IT resources and / or expertise to support analytical initiatives – 35%
- Increased urgency for fast information delivery and analysis – 29%

While many might automatically associate this type of flexible environment with open-source Hadoop-based technology, the research demonstrates that two thirds of companies are building their data lakes on proprietary, commercially available technology.

Figure 3: Insight with Oversight — Process and Technology Maturity



Source: Aberdeen Group, April 2017

Along with the sharing and distribution of data across business functions comes an elevated need for responsible oversight of that information. Best-in-Class companies are twice as likely as All Others to have strong data governance policies in place, supported by a technology backbone. Additionally, as companies try to take advantage of more data sources across the company, the traditional ad-hoc or batch-oriented data integration procedures may no longer be sufficient. Leaders are more than twice as likely to have real-time data integration capabilities, enabling them to make quicker tactical decisions supported by an increased variety of data.

Opening the Gates to the Data Lake

The increasing ubiquity of unstructured data, rich media files, and machine-generated Internet of Things (IoT) data underscores the need for flexibility in the data infrastructure. In response to this need, many companies are exploring a data lake that can handle a variety of information, comingled and managed in their native formats. While many might automatically associate this type of flexible environment with open-source Hadoop-based technology,

the research demonstrates that two thirds of companies are building their data lakes on proprietary, commercially available technology.

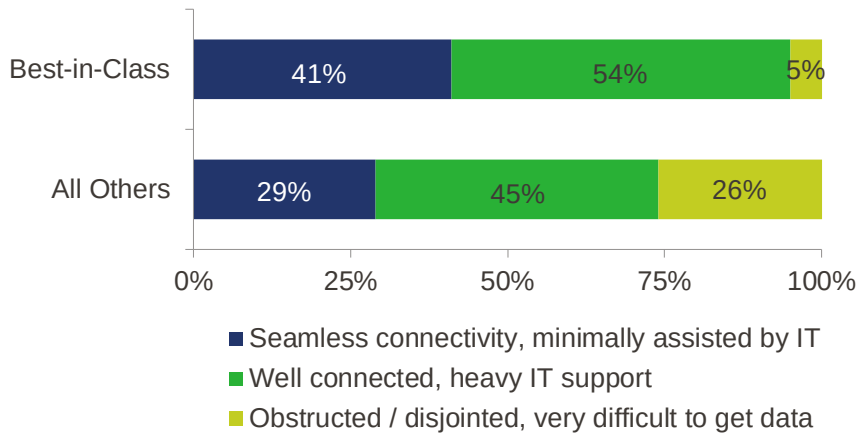
Aside from the technology foundation, arguably the most important aspect of a data lake is the ability to utilize the data stored within. The research shows that Best-in-Class companies are 27% more likely to have a data lake in place, but more importantly, it demonstrates that they have an easier time extracting information from it (Figure 4).

Fast Facts:

Top objectives for building / enhancing a data lake *(% of respondents)*

- Increase operational efficiency – 45%
- Make data available from departmental silos, mainframe, and legacy systems– 30%
- Provide a flexible data environment to support innovation – 30%

Figure 4: How Easily Can Users Connect with the Data Lake?



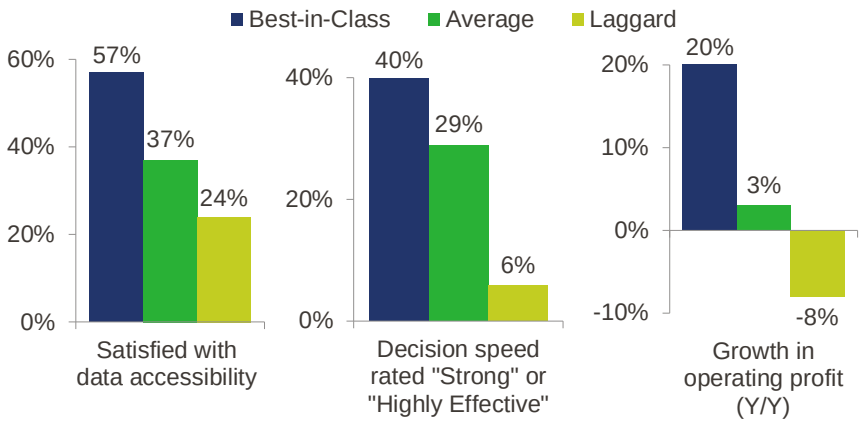
% of Respondents n = 342, Source: Aberdeen Group, April

While the creation and ongoing maintenance of a data lake will involve significant IT involvement, utilizing the contents shouldn't require a heavy technical hand. Most commonly, companies still require substantial IT support when accessing the data lake. However, Best-in-Class companies are more likely to report a seamless flow of information to business users. Average and Laggard companies, on the other hand, are more likely to experience significant impediments in their efforts to access the data lake.

Delivering Business Impact

Building and maintaining a reliable, flexible, and scalable data architecture is not an easy endeavor. Even in a successful environment, it's easy to get bogged down in the technical details of how systems are performing (e.g., petabytes available, data records accessible, sub-second query latency). The Best-in-Class, however, focus on transforming data into results. That process starts by enabling faster access to a wider variety of data (Figure 5).

Figure 5: Top Companies Transform Data into Results



n = 342

Source: Aberdeen Group, March 2017

With the ability to enrich their analyses with a broader array of information, and get the results they need in a timely way, Best-in-Class companies identify more opportunities for business growth and efficiency. By enhancing the beginning stages (or the back end) of the analytical process, top companies more easily recognize ways to diminish process waste, reduce cost, and ultimately deliver substantially greater improvement in operating profit compared with their peers.

9

Best-in-Class companies merge the foundational ideals of the data warehouse with newer capabilities to manage and move a diversity of data, positioning their organizations for better analytical efficiency.

Key Takeaways

Much in the same way that business analytics has become more of a philosophy for decision making than a suite of technologies, data management has followed suit in its ideology. With the data warehouse as a foundational element of their data strategy, top performing companies are investing their time and resources toward maximizing the value of their data. The following key takeaways summarize the most important findings from Aberdeen's research.

- **Data complexity has evolved from fear to expectation.** Growth in data volume is a challenge, but the bigger issue is the distributed, "siloed" nature of many environments today. There are simply more data sources, housing more data types, originating from more departments within the organization. Moreover, user expectations have increased to the point that non-traditional sources have become more commonplace in analysis.

- **The data warehouse isn't dead, just reborn.** The expansion in data volume and complexity has only underscored the need for efficient methods of data organization. Best-in-Class companies merge the foundational ideals of the data warehouse with newer capabilities to manage and move a diversity of data, positioning their organizations for better analytical efficiency.

- **Efficient data management is a strategic imperative.** Top performing companies are able to look beyond the challenge of managing data to the opportunity of exploiting it. Leveraging internal competency along with

technology Best-in-Class companies are able to explore more data, expedite the decision process, and execute on tangible business opportunities.

For more information on this or other research topics, please visit www.aberdeen.com.

Related Research

Data Preparation: The New Normal, The Now Necessary; March 2017

MDM without Limits: Connecting Data, Crafting Insight, Mastering Results; January 2017

Empowering Analytics: IT and Business Partnerships for Success; November 2016

The Hybrid Data Warehouse: Fluid, Flexible, and Formidable; May 2016

Author: Michael Lock, V.P. & Principal Analyst
Analytics & Business Intelligence

About Aberdeen Group

Since 1988, Aberdeen Group has published research that helps businesses worldwide improve their performance. Our analysts derive fact-based, vendor-agnostic insights from a proprietary analytical framework, which identifies Best-in-Class organizations from primary research conducted with industry practitioners. The resulting research content is used by hundreds of thousands of business professionals to drive smarter decision-making and improve business strategy. Aberdeen Group is headquartered in Waltham, MA.