# THE HYBRID DATA WAREHOUSE: FLUID, FLEXIBLE, AND FORMIDABLE

June 2016

➜ **Michael Lock,** Vice President & Principal Analyst,
Analytics & Business Intelligence

in t

## Report Highlights

| p2 | p5 | p7 | p8 |
|---|---|---|---|
| **A majority of IT and data professionals view data silos as their biggest challenge.** | **Top companies are almost twice as likely to use a hybrid data warehouse architecture.** | **Those using a hybrid data warehouse are 2.3 times more likely to be satisfied with information speed.** | **Hybrid data warehouse users saw almost twice the year-over-year increase in revenue as others.** |

As the foundation for most critical business decisions, today's data environments are not just a vital piece of IT infrastructure, but a key component of corporate strategy. However, as the user base has expanded and data has become more diverse, our methods for storing, managing, and organizing data must adapt. This report explores a new breed of data warehouse that can operate in a world of legacy on-premise systems while exploiting the potential of cutting edge technologies and deployment styles.
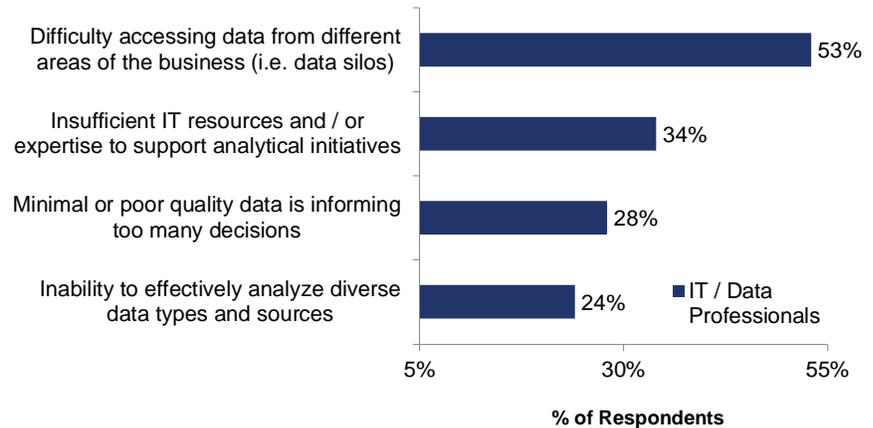
ABERDEEN GROUP

# 2

**The challenge for companies today, many beholden to legacy systems and IT bureaucracy, is that this growing demand for analytics often outstrips their ability service it.**

➜ Related Research, "Running Lean Analytics with a Cloud or Hybrid Approach"

➜ Related Research "The Visual Edge: Interactive Discovery vs. Traditional BI"

## Keeping up with Analytical Demand

It's no secret that analytics and data-driven decisions have become more prevalent in the business world. An influx of relevant data and a new breed of business-friendly and approachable tools have paved the way for more non-technical users to explore data and develop insights on their own. The challenge for companies today, many beholden to legacy systems and IT bureaucracy, is that this growing demand for analytics often outstrips their ability service it. Moreover, Aberdeen's research demonstrates that those IT and data professionals tasked with supporting an elevated level of analytical activity face a variety of daunting challenges when it comes to data (Figure 1).

### Figure 1: Key Challenges for Today's Data Professionals



Difficulty accessing data from different areas of the business (i.e. data silos) — 53%
Insufficient IT resources and / or expertise to support analytical initiatives — 34%
Minimal or poor quality data is informing too many decisions — 28%
Inability to effectively analyze diverse data types and sources — 24%

■ IT / Data Professionals

**% of Respondents**

n = 70, Source: Aberdeen Group, June 2016

While there is increased demand for analytics, the challenge doesn't necessarily relate to the ability to deliver more dashboards or more reports. The overarching challenge really lies in the foundation of the data used to fuel analytical activity. At the top of the list is the phenomenon of data silos. As more business users become active with analytical activity, their level of curiosity naturally rises. Why are sales slumping in the

# 3

northeast region? How are customers reacting to our recent product launch? These are valuable questions to ask, but to answer them properly they typically require data sourced from multiple areas of the company. Marketing data, finance data, sales data, product data, supplier data, etc. When data is locked away within these functional silos, the front-end analysis breaks down and users are left to make assumption-driven decisions rather than data-driven decisions.

Additionally, as new data sources come online there arises a two-pronged challenge. Companies are eager to exploit different forms of data (e.g. unstructured social media data, location-based data) but the infrastructure is often ill-equipped to house these disparate types and make them accessible. Moreover, along with new data sources comes human error and issues with quality and usability of data. Looming large over all of these issues is the fundamental lack of IT manpower and experience to tackle all this data complexity, as well as the ability to deliver the type of analytical experience companies are striving for these days.

## Best-in-Class Data Management

In light of these challenges, top companies are taking action in a variety of areas. Aberdeen's recent data preparation report demonstrated several defining characteristics of Best-in-Class companies. From an organizational standpoint, they look to make their data more accessible and develop analytical skills internally (see sidebar). From a technology standpoint, these leading companies look to build out capabilities in three main areas:

- **Rapid retrieval.** A major challenge that affects decision makers today is data urgency. Operational and strategic decision makers alike struggle to get the information they

**Fast Facts**

Best-in-Class companies are:

- 88% more likely share data (securely) across business functions

- 45% more likely to have data governance policies in place

- 68% more likely to develop analytical skills in-house

As compared to all other companies

# 4

**This disparity of technology and infrastructure is largely what drives the need for a hybrid architecture, one that can be flexible in serving multiple types of analytical tools, as well as deployment options.**

need, when they need it. Timeliness starts in the data infrastructure and Best-in-Class companies are leveraging cutting-edge data warehouse technology that allows for sophisticated querying and the rapid retrieval of data from large and complex data environments.

- **Flexible deployment.** Most analytically-inclined companies, particularly the larger ones, have been making technology investments for years. Investments have been made in a variety of data processing engines and analytical tools but also in terms of the deployment approach (e.g. SaaS, public cloud, private cloud, or traditional on-premise solutions). This disparity of technology and infrastructure is largely what drives the need for a hybrid architecture, one that can be flexible in serving multiple types of analytical tools, as well as deployment options.

- **Scalability.** Data growth is nothing new, but the explosion in multiple varieties of data (e.g. unstructured data, machine / sensor generated data, location-based data) has underscored the need for a scalable data environment. Many companies address the need for scalability and flexibility simultaneously with the use of open-source technologies like Apache Hadoop that allows for large scale management of a variety of different data types.

The research shows that Best-in-Class companies are leveraging technologies to improve these three areas (Figure 2).

**Figure 2: The Right Technologies, The Right Backdrop**

5



**% of Respondents**                    n = 174
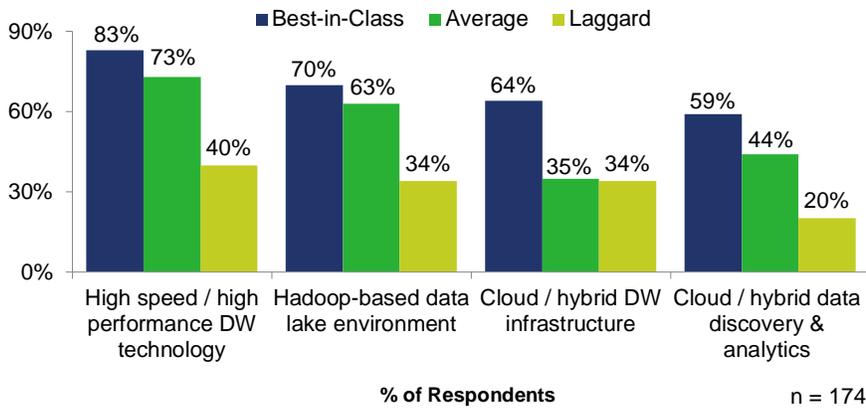
Source: Aberdeen Group, June 2016

**Fast Facts**

Best-in-Class activities / technologies in use for data preparation *(% of Respondents)*

- Profiling – 79%
- Cleansing – 64%
- Transformation – 57%
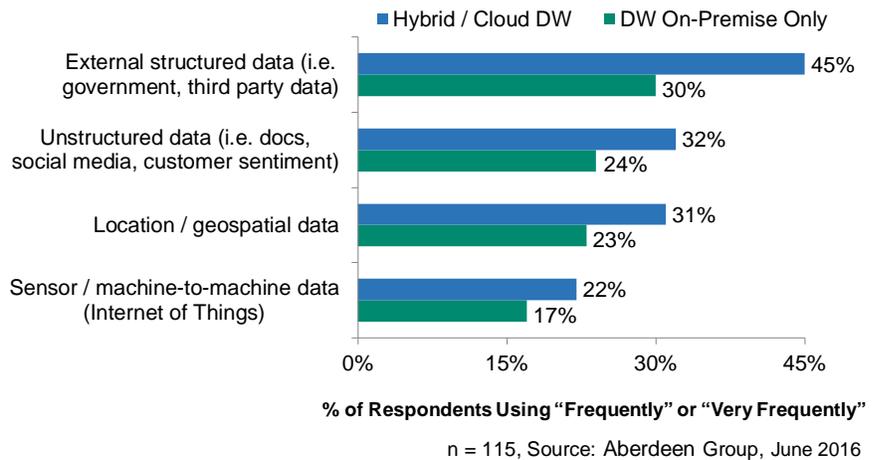- Modeling – 57%
- De-duplication – 43%
- Curation – 43%

One thing important to note is the Best-in-Class focus on hybridity, both in their data infrastructure and with their analytical tools as well. Given the substantial investments, companies have made in their IT environment, a hybrid approach allows them to utilize these investments to the best of their ability while exploring more flexible and scalable cloud-based solutions as well.

## Characteristics of the Modern Data Warehouse

Following on this concept of a hybrid data warehouse architecture, the research demonstrates a variety of noteworthy characteristics that these companies share. First, those with this type of modern data warehouse approach are able to exploit not only the growing volume of data on hand, but also the expansion of data types as well (Figure 3).

ABERDEEN
GROUP

6

**Figure 3: A Breadth of Data in Use**



% of Respondents Using "Frequently" or "Very Frequently"

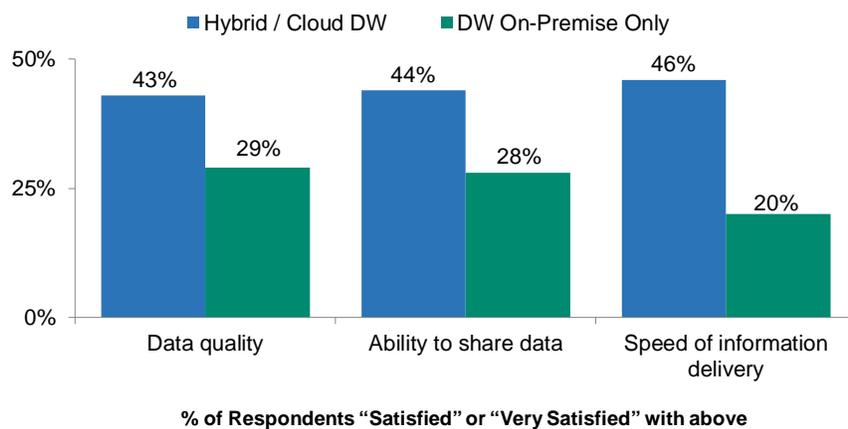n = 115, Source: Aberdeen Group, June 2016

Traditional structured data from business applications and other transactional systems still comprises the bulk of information used by companies today, but most are looking for ways to enrich their analyses with different forms of data such as unstructured data from documents or social media outlets and location-based, geospatial information. Combining traditional data with these other forms of information can provide important correlations to inform critical decisions. For example, combining customer sentiment data from social medial channels with internally structured product data can reveal insights to help shape product roadmaps and market messaging. Supply chain analyses enriched with geospatial data can uncover opportunities for greater operational efficiency and cost reduction. Companies taking a hybrid data warehouse approach are able to exploit these opportunities for enriched analyses and new insight.

Another important characteristic of these companies with a modern data warehouse environment is their ability to understand and deliver against users' needs for fast and fluid data. The research shows that these companies are more likely

7

to have processes in place to gather and deliver on user requirements for analytics. These companies are also more likely to hire and develop analytical talent capable of exploiting their growing volumes of data. Perhaps more importantly, though, these companies are more likely to report user satisfaction in several key areas of the data environment (Figure 4).

### Figure 4: Fast and Fluid Data



■ Hybrid / Cloud DW    ■ DW On-Premise Only

**% of Respondents "Satisfied" or "Very Satisfied" with above**

n = 115, Source: Aberdeen Group, June 2016

Data fluidity is an interesting concept that is top-of-mind for many data professionals these days. Fluid is synonymous with smooth, effortless, and graceful. At its core, fluid data implies a lack of (or minimal) barriers to the movement and exchange of data. Poor data quality is a barrier because it causes wasted time and effort in the analytical process. Data silos act as a barrier to the exchange of ideas and insight across the business. Ultimately, all of these barriers produce a slower propagation of data in the company and lead to late delivery of information and missed opportunities. Companies with a modern data warehouse approach are able to improve these three key areas of the data environment, reduce barriers to information flow, and empower their users in the process.

### Fast Facts

Those with a cloud or hybrid data warehouse architecture are:

- **3.8 times** more likely to use cloud-based traditional BI / managed reporting
- **4.3 times** more likely to use cloud-based data discovery / interactive visualization
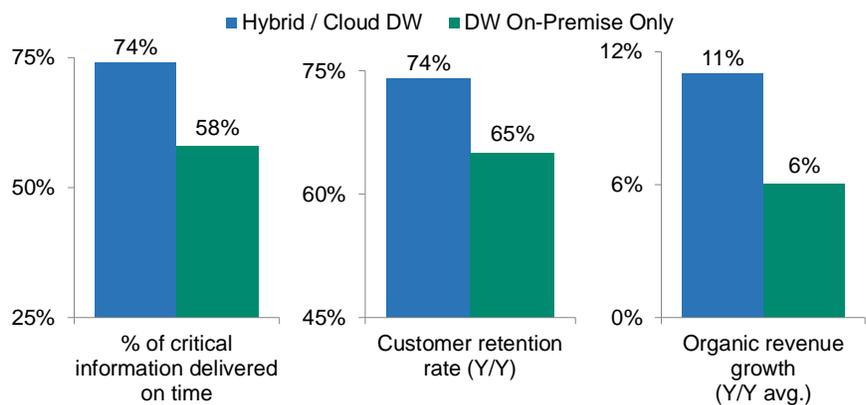- **3.2 times** more likely to use cloud-based predictive analytics

8

**An efficient data environment coupled with strong analytical activity leads to opportunities for customer cross-sell and upsell, higher retention, and higher revenue.**

## Driving Business Outcomes

Ultimately these efforts are all for naught without tangible results. The research demonstrates that these organizations, in addition to having greater organizational maturity and breadth of technology usage, also experienced elevated performance against three key metrics (Figure 5).

**Figure 5: Data Transformed into Results**



Legend: ■ Hybrid / Cloud DW   ■ DW On-Premise Only

% of critical information delivered on time: 74% vs 58%
Customer retention rate (Y/Y): 74% vs 65%
Organic revenue growth (Y/Y avg.): 11% vs 6%

n = 115, Source: Aberdeen Group, June 2016

Their focus on enabling data fluidity allows for faster and cleaner exchange of information, producing more timely delivery of key insights. These insights then help organizations deliver against the urgent demands of their customer base, helping to improve customer satisfaction and retention. A variety of factors can contribute to revenue growth but perhaps none more directly than a close understanding of existing customers. An efficient data environment coupled with strong analytical activity leads to opportunities for customer cross-sell and upsell, higher retention, and higher revenue.

## Key Takeaways

If top-notch data-driven decisions and game changing insights form the pinnacle of analytics for today's companies, a strong

ABERDEEN GROUP

9

underpinning of clean, reliable, and timely data is the foundation. Given the evolution of analytical technologies and the supporting infrastructure, it's no surprise that many companies now face a complicated web of legacy hardware and software systems loosely coupled with more modern cloud-based architectures. Those taking a hybrid approach in their data warehouse environment are finding themselves well-equipped to manage their data and infrastructure challenges. According to the research findings, three main concepts are worthy of highlighting:

➔ **Data silos are a major problem (and an opportunity).** The thirst for richer and more diverse data to support analyses is prevalent among more users today, both technically-inclined and not. It's therefore not surprising to see that the data silo effect is the top data-related challenge impacting companies these days. On the other hand, companies that can effectively address this challenge and improve the accessibility and portability of data have a lot to gain. Top companies put measures in place to govern and control the usage of data while at the same time improving the flow of information across the company. These leading organizations enjoy a higher degree of data fluidity and user satisfaction as well.

➔ **Top companies explore cutting-edge data warehouse technology.** Similar to the data silo effect discussed above, the challenge of complexity and disparity in the data environment is a double-edged sword. There is an obvious challenge of managing and extracting value from antiquated systems while exploring new technology, but there is inherent opportunity as well. Companies utilizing a hybrid data warehouse environment in conjunction with cutting edge technology for speed and scalability are

more likely to navigate these challenges gracefully and make better use of their data.

➔ **The results are enticing and attainable.** It's never simple to put together a watertight ROI case for IT implementation. However, these days, the impact of effective analytics is more easily translated to upper management than ever before. A strong and flexible data environment leads to efficiency in the decision process which generates richer and more timely insights. Those data-driven insights then lead to a variety of opportunities for business growth and efficiency, ultimately having a tangible impact on the income statement.

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

| Related Research |
|---|
| *Running Lean Analytics with a Cloud or Hybrid Approach*; April 2016 | *A BI Blessing from Above: The Impact of the Executive Touch*; March 2016 |
| *The Visual Edge: Interactive Discovery vs. Traditional BI*; April 2016 | *Three Levels of ROI from Data Quality Initiatives*; March 2016 |
| *Data Preparation and the Evolution of Analytics*; March 2016 | *BI Excels in the Mid-Market: A Nimbler Version of the Enterprise*; February 2016 |
| Author: Michael Lock, Vice President & Principal Analyst, Analytics & Business Intelligence (michael.lock@aberdeen.com) | |

11

**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 Boston, MA.