

December 2012

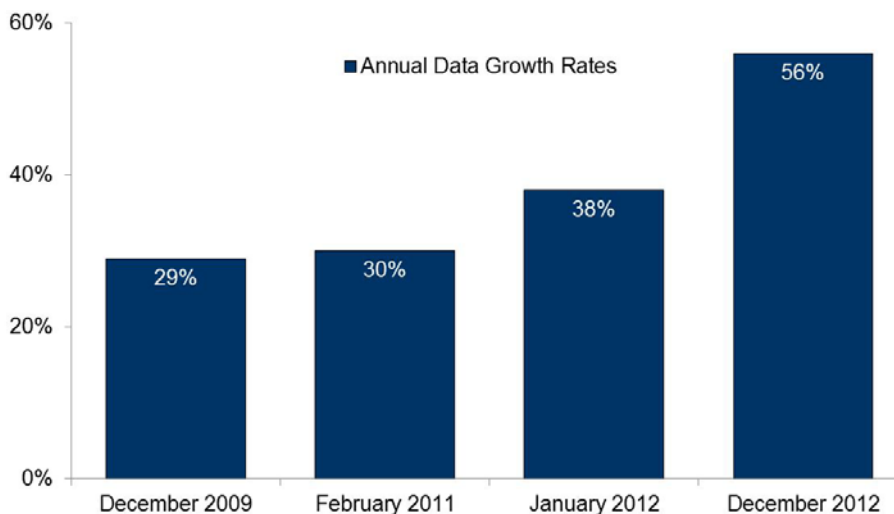
## The Big Data Imperative: Why Information Governance Must be Addressed Now

Aberdeen has long illustrated the benefits of well-managed, trustworthy data, and the problems associated with poor data quality. As data volumes rapidly expand and data environments become more complex, what were once small nuisances evolve into massive, company-wide problems. Using research collected from December 2009 to December 2012, Aberdeen will examine the hidden financial penalties for lapsed information governance and untrustworthy data. In order to avoid these pitfalls and achieve better business efficiency and operational performance, top performing organizations have found it necessary to invest in tools such as data security, master data management, data quality, data lifecycle management and data integration.

### Why Information Governance?

To put it simply, the amount of business data is growing at a faster rate than any previous time in history. Aberdeen's December 2012 [Data Management for BI](#) research puts the current growth rate at 56% year-over-year. As Figure 1 displays below, this growth rate has quickly increased over the past four years, and shows no sign of stopping. Not only are the storage requirements doubling every 19 months, but all the associated costs for storing, managing, and protecting this data are exploding as well.

Figure 1: The Exponential Explosion of Business Data



Source: Aberdeen Group, December 2009 ~ December 2012

### Research Brief

Aberdeen's Research Briefs provide a detailed exploration of a key finding from a primary research study, including key performance indicators, Best-in-Class insight, and vendor insight.

### Definition

✓ **Information Governance** refers to the combination of people, processes, and technology that ensure business data is trusted, secure, and able to be leveraged as an enterprise asset. Critical technology elements include data integration, data quality, master data management, data lifecycle management, and data security.

Every day, new data sources spring up to provide additional insight into economic trends and customer behavior. These new sources provide valuable opportunities for organizations to optimize their products and operations, but they can come with a cost. On average, Aberdeen's research shows that the more data sources a company uses, the lower the trust in their data becomes. Among organizations that reported fewer than 20 unique data sources that were routinely accessed and analyzed, 70% had high levels of trust in their data (ranked as seven or higher on a scale of one to ten). For companies that reported 20 or more unique sources, this rapidly dropped to 43% that trusted their information. This means that before organizations can take full advantage of the opportunities these new data sources provide, they must establish a solid foundation of data management and trust to build upon.

As covered in the following sections, companies that are not taking measures to ensure business data is accurate, protected and easily managed at large scale, are losing millions of dollars in avoidable costs. They are wasting valuable resources on inefficient processes, suffering lower product sales, and losing more customers to their competition. These financial penalties are severe now, but could easily become crippling in the near future. If we assume that the growth of data remains static at 56% (a very conservative estimate), five years from now the volume of data and all associated costs for supporting it will be almost 6-times larger than today. In other words, while the current data problems might seem onerous and expensive to fix, the cost and difficulty will be exponentially greater in the future.

When addressing these issues through data management solutions or a Big Data initiative, it is important to properly manage unstructured data (see sidebar). Most companies, however, begin their journey by addressing their traditional, structured information. Over 95% of companies with a Big Data initiative said their structured data was critical to their analytic success, which indicates how important information governance solutions are. In brief, information governance covers a range of capabilities that ensure data from all across an organization is accurate, standardized, easily accessible, and protected. Some of the major elements of information governance are listed below:

- **Data integration.** Being able to move data from one system, application, or database to another is a crucial component to providing data access throughout an entire organization. With the number of data sources growing, and the complexity inherent in having multiple data formats and systems, these integration tools are becoming more important than ever. Recent trends in integration are explored in [Beyond Agile Analytics: Is Agile Data Integration Next?](#) (June 2012).
- **Data quality.** These tools work to make all business records comply with industry and company standards, are stored in the

#### Definitions

- √ **Big Data** refers to the problems of capturing, storing, managing, and analyzing massive amounts of various types of data. Most commonly this refers to terabytes or petabytes of data, stored in multiple formats, from different internal and external sources, with strict demands for speed and complexity of analysis.
- √ **Unstructured data** refers to data stored in files, documents, presentations, spreadsheets, web pages, email messages, instant messages, images, audio files, video files, etc. While each of these formats do indeed have "structure," conventional use of the term unstructured data is intended to distinguish from data stored in *structured* formats (e.g. in databases).

proper format for use by software applications, and present the most up-to-date version of information. A closer look at the value of data quality tools can be found in [Data Quality and the Supply Chain](#) (October 2011).

- **Master data management (MDM).** This technology provides a centralized repository for critical business data, like customer lists or product information. This repository serves as the authorized "single version of the truth," and often feeds other important enterprise applications like ERP and CRM systems. For more information, see [The State of Master Data Management, 2012: Building the Foundation for a Better Enterprise](#) (May 2012).
- **Data lifecycle management.** These tools help manage data from initial requirements until final retirement. They improve IT efficiency, agility, and risk mitigation by optimizing retention and right-time access to the right data. A critical component of these solutions is identifying when information is no longer required for day to day business operations so it can be removed from production systems and moved to less expensive storage media while still providing access. See [Archive Agility: Building Business Resilience through Active Archiving](#) (July 2011) for more information.
- **Data security.** This covers a wide range of different tools that control which employees have permission to access what information, enforce those access rules, and ensure personally identifiable or valuable information is securely stored behind the corporate firewall. A deeper dive into several important facets of data security is presented in [DLP, the Ideal Referee: Let the Game Go On!](#) (November 2011).

## The Cost of Doing Nothing

When economic recession hits, the natural reaction is to tighten the purse strings and eliminate all non-essential spending. While this is certainly suitable for weathering a depression in the short term, at some point organizations have to renew their investment in improving operations and remaining competitive. The temptation is to minimize the impact on the organization's bottom line by maintaining the status quo, but Aberdeen's research shows how costly and detrimental this stagnation can be.

### Unreliable Business Data, Poor Business Decisions

Information governance solutions focus entirely upon improving business data, so it comes as no surprise that organizations without these tools report having worse data quality than companies that made the investment. Aberdeen's [Master Data Management](#) (May 2012) research shows that all companies wanted their business records to be more than 90% complete and accurate, which includes having correct information in all essential and most supplemental fields. The companies that chose not to invest in data quality and MDM tools reported the actual state of their data as only 70%

"When addressing the factors that contribute to high quality data, try to follow these steps:

- √ Define
- √ Measure
- √ Analyze
- √ Improve
- √ Control

Set clear accountabilities for data quality. Measure the performance of Senior Management through the use of data quality target metrics. Create data quality awareness throughout the entire company by launching a formal training program."

~ Director of Human Resources  
Large (>\$1B) European Software company

accurate, a far cry from their goal and 12% worse than companies with governance tools.

The impact of this higher-quality data is most immediately evident in the quality and speed of business decisions. Organizations reporting data accuracy of over 90% were able to put reliable information in the hands of their executives fast enough to meet demand four times out of five (Figure 2). Companies with lower than 70% data accuracy only succeeded at meeting this demand window half of the time. Not only was the data delivered on time, its accuracy meant the resulting decisions were more likely to be correct. Seventy-nine percent (79%) of companies with high-quality data rated their decision making as a 7 or higher on a scale of one to ten. Companies with less than 70% data accuracy were unhappy with their decisions more than two thirds of the time.

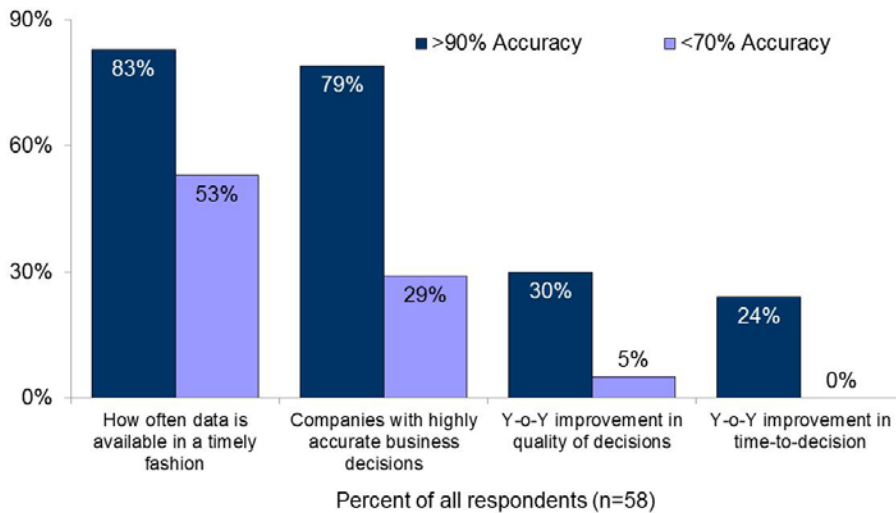
"When it comes to data management and Big Data programs, just do it. It may be an expensive change process, but so is lost opportunities and lost business due to poor data quality."

~ CEO

Small IT Services Company

EMEA

**Figure 2: Accurate Data Equals Better Business Decisions**



Source: Aberdeen Group, December 2012

Furthermore, companies with higher levels of data accuracy improved their performance at a faster year-over-year rate. Both the quality and the speed of their decisions saw more than a 24% improvement in the last fiscal year, while companies with low data quality remained more or less static. The importance of timely, accurate business decisions cannot be stressed enough. Careers are made or lost and companies succeed or fail based on how executives set goals and react to market changes. Outperforming other organizations in this area is a powerful incentive for investment; in fact, 56% of the Best-in-Class companies in Aberdeen's [Data Management for BI](#) (December 2012) research reported that using faster and better analytics to gain a competitive advantage was their top goal for a Big Data program.

Given that data is growing so rapidly, this data quality problem is getting increasingly more difficult to manage. More importantly, it causes the gap between companies taking active steps to fix their data and companies that

"Data quality reflects the level of responsibility that the business demands and delivers; if no one cares then data suffers."

~ Manager

Large (>\$1B) European Oil / Gas company

do nothing to grow even larger. Table I shows that organizations with information governance tools (including data quality and MDM) not only have more accurate data to start with, but they are improving at almost 3-times the rate of their competitors. With more reliable data, they in turn see improvements in business efficiency. Employees are able to quickly find the data they need without searching through multiple systems or worrying if they have the latest version. Every data-centric task, from billing and invoicing, to customer interactions and product development, benefits from faster data access.

Companies that do nothing, on the other hand, see their competitor's performance and the overall growth of their data rapidly outpace any moderate gains.

**Table I: Addressing Data Quality**

Performance Metric (year-over-year change)	With Data Governance	Without Data Governance	Cost of Doing Nothing
Data accuracy	21% improvement	8% improvement	<b>2.6-times slower improvement</b>
Data completeness	23% improvement	8% improvement	<b>2.9-times slower improvement</b>
Time-to-information	26% improvement	10% improvement	<b>2.6-times slower improvement</b>
Time-to-decision	18% improvement	6% improvement	<b>3-times slower improvement</b>

Source: Aberdeen Group, December 2012

### **No Business Agility and More Wasted Time**

Solutions to improve the efficiency of a data center or streamline the IT architecture tend not to have a dramatic impact on a company's revenue numbers. However, even small improvements in the effective use of resources can have a trickle-down effect on all business units. One such example is using the data integration element of information governance. In particular, extract, transform, and load (ETL) tools can easily move data from location to location and rapidly connect to new data sources - and the valuable insights they can provide. This helps companies react faster to business pressures by accessing new information and spend less time on these essential tasks.

In Aberdeen's [Data Integration](#) (June 2012) research, companies that had done nothing to improve their data integration capabilities and found themselves without these ETL tools reported an average of 61 days to integrate a new data source into their systems. This was 33% slower than companies that had adopted ETL solutions, which accomplished the same task in only 41 days.

These new data sources are often the result of a merger or acquisition, when the systems of two different organizations must be combined. However, they are also an integral part of Business Analytics initiatives, as companies frequently supplement their own records with information that adds perspective on their customers, product performance, or economic trends. This is especially true in Big Data environments, which on average include 20 major internal sources (like databases and software applications), 12 sources from business partners, and nine other external sources such as economic, weather, or census data. This total of 41 unique data sources is 20% more than non-Big Data companies, and represents a significant amount of time and effort by a dedicated IT team.

To illustrate the impact of this technology, a Big Data company without ETL tools spends 20 extra days to integrate each data source. Extrapolating this across the 41 different data sources an average company uses means they would spend an extra 820 days working on these projects. That is over 2.3 years slower than their competitors; valuable time that could be spent on other critical initiatives, and time during which business units are unable to access this new information.

Due to the difficulties of navigating multiple databases and systems and not having access to valuable data sources, business units not supported by integration tools also reported that the time spent on their daily data-centric processes did not improve over the last year. In fact, they had a 1% decrease in performance. Companies with integration tools, however, were able to streamline all these operational tasks by an average of 15% in the last year. That is 15% less time spent on tasks such as processing invoices, responding to customer requests, managing inventory, evaluating loan applications and submitting insurance claims. Again, organizations without these critical solutions waste more time, while watching their competitors become lean and efficient.

### **Unable to Manage Rampant Data Growth**

Another technology with similar business impact is data lifecycle management. A critical component of these solutions is identifying when information can be moved to cheaper, long-term archive storage, while still keeping this data accessible for emergencies. The first step in this process is knowing what applications or data sources have the highest priority and are accessed the most frequently, and which ones are less relevant for the business. The costs of keeping an application or a large volume of data online include administration, network bandwidth, power, space, and storage hardware; archiving the same information requires only a fraction of the cost.

One common use case is during a merger or acquisition, when the two companies each have a version of an enterprise application like customer relationship management (CRM) or enterprise resource planning (ERP). It is not efficient or cost effective to keep both versions running, yet it is also practical to store the obsolete version in case the information or data rules

#### **Fast Facts**

√ Organizations using data lifecycle management tools reported that **53%** of their archived information was tagged with metadata for easy searching and access

Compared to

√ **29%** of archived data for companies without these tools

are needed in the future. Almost two thirds (64%) of organizations that used data lifecycle management tools were able to reduce costs through archiving such applications while still maintaining live access to the application data. Only 36% of organizations without these tools could accomplish the same task.

Another cost-reducing technique is called Tiered Storage, which involves putting the most frequently accessed information on fast, expensive storage devices like solid state disks (SSD) or fibre channel drives. When this information ages and becomes less relevant, it is automatically moved to slower, cheaper drives like serial ATA (SATA) or serial attached SCSI (SAS), and finally retired to an archive. This simple use of data lifecycle management tools can minimize an organization's spend on storage hardware, while still supporting fast data access for the line-of-business.

**Table 2: Managing Data Storage Costs**

Storage Tiers	With Lifecycle Management	Without Lifecycle Management
<b>Amount of data in Tier 1 storage (SSD, Fibre Channel)</b>	12% 29 TB	25% 60 TB
<b>Total cost for Tier 1 storage (\$3,750 per TB)</b>	\$108,750	\$225,000
<b>Amount of data in Tier 2 storage (SATA, SAS drives)</b>		
<b>Amount of data in Tier 2 storage (SATA, SAS drives)</b>	21% 50 TB	24% 55 TB
<b>Total cost for Tier 2 storage (\$2,830 per TB)</b>	\$141,500	\$155,650
<b>Amount of archived data (Tape storage)</b>		
<b>Amount of archived data (Tape storage)</b>	67% 161 TB	52% 125 TB
<b>Total cost for archive storage (\$200 per TB)</b>	\$32,200	\$25,000
<b>Total cost for data storage</b>	\$282,450	\$405,650
<b>The cost of doing nothing this year:</b>	<b>\$123,200</b>	
<b>The annual cost in five years:</b>	<b>\$729,600</b>	

Source: Aberdeen Group, July 2011

According to Aberdeen's July 2011 [Archiving](#) research, organizations with data lifecycle management tools stored less data on the expensive drives, kept it there for a shorter period of time, and successfully archived a larger percentage of their business information without disrupting operations. The cost differences between different storage tiers can be striking; Aberdeen's

**Maturity Class Framework**

In the July 2011 [Archiving](#) report, companies were categorized based on their performance in the following metrics:

**Best-in-Class (Top 20%)**

- √ 20 minutes to recover an archived file
- √ .5 business interruptions per year
- √ 36 minutes for longest period of downtime in 12 months

**Industry Average (middle 50%)**

- √ 12.4 hours to recover an archived file
- √ 2.1 business interruptions per year
- √ 5.7 hours for longest period of downtime in 12 months

**Laggard (bottom 30%)**

- √ 44 hours to recover an archived file
- √ 3.3 business interruptions per year
- √ 11 hours for longest period of downtime in 12 months

report on [The Ultimate Storage Tier](#) (May 2012) revealed the average cost for top tier storage was \$3,750 USD per terabyte. Archive storage was a mere 5% of this amount, or \$200 per TB. Given that an average organization in this research study stored 240 TB of business data, such a company could save over \$123,000 dollars a year through using data lifecycle management tools (Table 2). While this might seem like a manageable cost, remember that data doubles every 19 months. If a company ignores this issue for five years, and the growth rate remains static, this annual cost would balloon to \$729,000 per year, and the total cost during this time would reach over \$1.8 million. Again, this analysis only covers the hardware involved. The fully loaded costs of storing this data include labor, datacenter space, power and cooling, and can end up being many times greater than the numbers cited above.

A second use-case for data lifecycle management is for application development and testing. With the fast-paced, iterative process required to quickly create new enterprise software, it becomes imperative for organizations to have effective and efficient testing strategies which match the same pace and rigor of live environments — complete with processes for managing cloned versions of live data. Data lifecycle management tools allow these developers to select viable live data to use and manage it throughout the development process while still remaining compliant with regulations around data privacy and access to sensitive information. This is an often overlooked aspect of data security, as cloned data is not always accounted for in corporate security policies, but can easily cause a data-exposure event. As the following section will detail, the seriousness of such incidents should not be taken lightly.

Finally, organizations need the tools to quickly and efficiently manage data within their archive. In Aberdeen's July 2011 research into [Archiving and Business Resiliency](#), 75% of Best-in-Class companies reported that they had implemented archive management software tools to help move data to an archive and understand how long legal mandates required that it be kept. Only 20% of Laggard organizations possessed these tools.

For the Laggards, the absence of these solutions directly correlated to their performance in retrieving archived information on demand. They reported an average time of 44 hours to recover data from the archive, which was a startling 132-times slower than Best-in-Class, who only needed 20 minutes. For the 46% of large companies that need to recover archived data at least once per day, the impact of not having these lifecycle management tools equals 11,000 hours per year spent waiting. Not only must data be kept for mandated periods of time, but it has to be accessed and produced for legal and audit purposes. The ability to quickly conform to these demands can easily reduce the complexity of these proceedings and eliminate fines or penalties for not being able to turn over the information in a timely fashion.

#### Definitions: Company Size

Aberdeen defines company size based on the following categories of annual revenue:

- ✓ **Large companies** report annual revenue greater than \$1 billion
- ✓ **Mid-sized companies** report revenue between \$50 million and \$1 billion
- ✓ **Small companies** report revenue under \$50 million



### The Hidden Costs of Data Security

Even if a company keeps their IT environment static, the rest of the world moves on. From a data perspective, this means improved tactics from the criminal element, new viruses and malware, new software vulnerabilities to be exploited, and new industry and governmental regulations on data privacy. It can often be hard to justify IT expense to combat these risks. After all, the mark of a successful security implementation is that... nothing happens. Quantifying the "what-ifs" and "might-have-beens" can be difficult, but Aberdeen's [Data Loss Prevention](#) (November 2011) report revealed that the average annual costs absorbed by companies as a result of failing to properly secure their business data can easily reach millions of dollars (Table 3).

"My number one comment or words of wisdom for data loss prevention projects, based on experiences with numerous clients, is actually not a technical one. You must educate your employees about the huge consequences and liability in the event of data loss!"

~ Executive, VP, IT  
IT consulting service  
Asia/Pacific region

**Table 3: Annual Costs of Data Security Events**

Category	# of audit events per year	Average cost of audit events	# of data loss events per year	Average cost of data loss events	Total annual cost (adjusted per TB)
Companies without data security tools	6.3	\$24,000	2.3	\$130,000	\$449,000 / TB
Companies with both passive and active data security tools	1.7		1.4		\$223,000 / TB
<b>Annual cost of not implementing data security tools</b>					\$226,000 / TB
<b>Total annual cost of doing nothing: (for an average Big Data organization with 180 TB of business data)</b>					<b>\$40+ million dollars</b>

Source: Aberdeen Group, December 2012

Audit or noncompliance events result from failing to meet regulatory standards, such as the proper procedure for encrypting, obfuscating and restricting access to credit card numbers, medical records or personally identifiable information (PII). The financial impact for failing can include fines, increased oversight, and the cost of remediating the problem. Data loss or exposure events, on the other hand, occur when sensitive information falls into hands outside the organization through malicious activity like corporate espionage, or negligence like leaving an unprotected laptop in a taxi. The financial impact here can be staggering; remediation costs are compounded with legal settlements to wronged customers and lost sales due to negative press and damage to the corporate brand. Having an incident broadcast on national news can have further, incalculable costs in terms of lost customers and reduced loyalty. After all, if customers can't trust a company to keep their sensitive data safe, what impetus do they have to continue bringing them their business?

As Table 3 shows, companies were able to reduce the number of both types of events through passive and active data security tools. Passive tools involve monitoring and flagging unauthorized events, while active tools take action to block the action or quarantine the file. Companies that did not make the investment in these security tools reported 3.7-times more events per year. When calculating the total cost, and adjusting it for dollars spent per terabyte (TB) of data, they ended up spending \$226,000 more than their competitors. Given that the average Big Data organization stores 180 TB of data, this extrapolates to a hidden cost of over \$40 million dollars per year. This reinforces how important data security is in large-scale data environments. While a smaller company might only have data on a few hundred customers at risk, larger enterprises often carry hundreds of thousands or even millions of sensitive records, making a single major breach that much more expensive.

### **The Benefits of Treating Information as an Asset**

Doing nothing to improve a data environment can result in significant hidden costs and financial penalty. However, the flip side of the coin is that operational benefits are available to companies that take the time and spend the resources to treat information as an important asset. The following are several ways organizations use information governance tools to achieve performance improvements in critical areas.

#### **Achieving the 360-Degree View of the Customer**

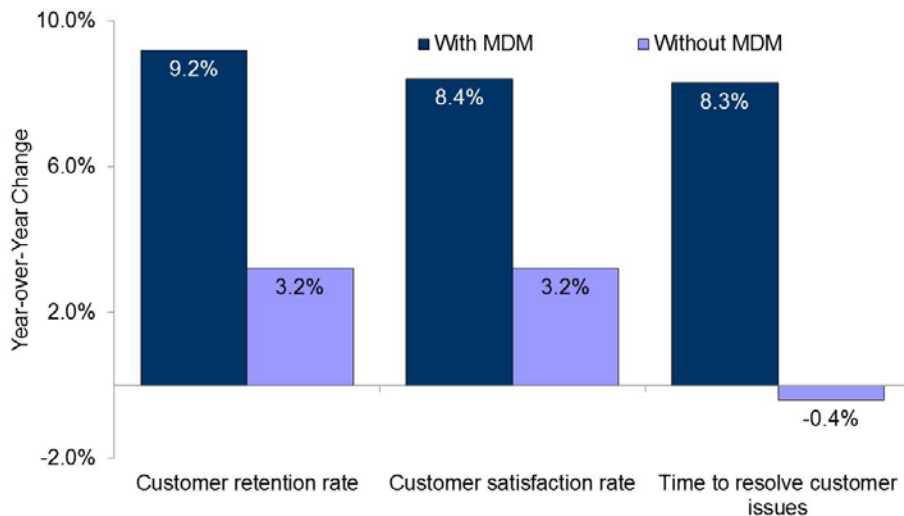
Data, such as purchase history, demographics, or activity on a company's website and online store, forms the foundation for many customer interactions. As such, companies are always seeking new ways to collect and analyze their customer data, in order to better target their products and deliver advertisements or sales opportunities to the right people at the right time. This ability to completely understand one's potential buyers is often referred to as "the 360-degree view of the customer." Aberdeen's February 2012 [Enterprise Resource Planning](#) research highlighted how certain information governance tools, specifically master data management, played an important role in improving this customer insight (Figure 3).

Master data management is an ideal fit for customer analytics, as it can collect disparate elements of the overall picture and combine them into a single, centralized, cohesive unit. This authoritative data source can then in turn feed different enterprise applications and business units. That way when the company's online store tries to determine which items to highlight during a shopping session, or their customer service center receives a complaint call, or their sales team tries to renew a service agreement, they all work from the same up-to-date records. The companies that implemented an MDM system were able to report significant performance improvements compared to companies without. As Figure 3 shows, over a 12 month period they were able to keep more of their customers happy, keep them coming back for more, and resolve complaint tickets faster.

"Understand the requirements before you start an MDM initiative. Due to the amount of effort and money vested, reworking is usually not an option. Short-sighted decisions typically can't be reversed except at the very beginning, and you might end up living with a sub-standard result until the business pressure forces you to do it all again."

~ Manager  
Large (>\$1B) North American  
Paper / Lumber / Timber  
company

**Figure 3: Clean Data Equals Better Customer Analytics**



\*Positive result indicates improvement, negative result indicates performance decrease

Source: Aberdeen Group, December 2012

With better touch points with their customers, it is no wonder that these companies also reported a year-over-year increase of 15% in overall product sales and a 10% increase in profitability.

### **The Rewards of Trustworthy Data**

A profound, yet often overlooked, benefit of good information governance is simply that employees place more trust in the data. This little change can have a widespread impact throughout the organization, as employees no longer need to constantly double check their information and decision makers are able to rely on concrete facts instead of external analysis and conjecture. Building this type of data-driven business culture takes time, dedication, and the right governance tools. The May 2012 [Master Data Management](#) research asked organizations to rate their level of trust in their business data on a scale of one to ten. Among companies showing High or Very High levels of trust (seven or higher on the scale), 77% had implemented information governance tools, which was more than twice the adoption rate of companies with lower levels of trust. Additionally, 49% of the high-trust companies reported having these solutions for five or more years, indicating that they took the time to gradually build up the quality of their systems and improve the interaction between employees and data.

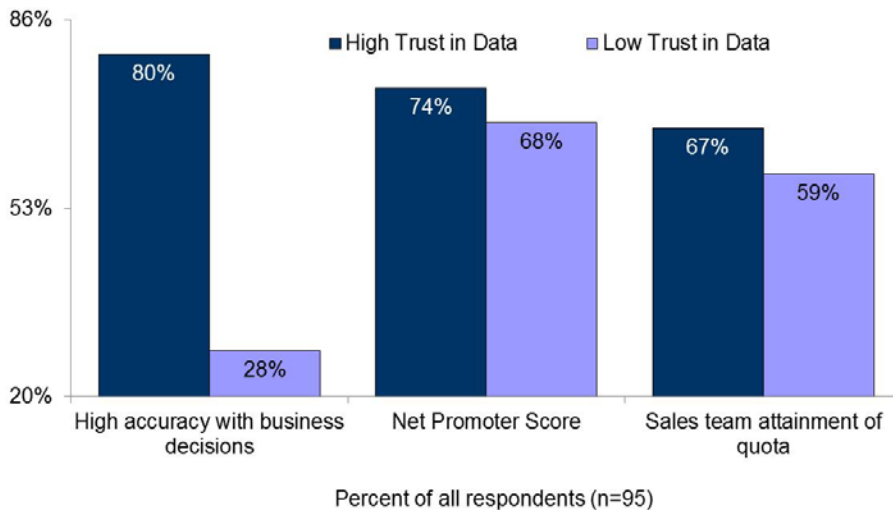
Having information governance tools leads to trustworthy data, and trustworthy data leads to better business performance. Organizations reporting high levels of trust in their information also enjoyed more accurate business decisions which, like the MDM tools above, specifically contributed to better understanding customers and meeting their demands (Figure 4). The sales teams in these companies had a greater degree of success interacting with potential buyers. More than two-thirds (67%) of

"It is always more important, according to upper management, to produce cost savings and deliver on the initiatives that the 'business' sets and understands, than it is to support and improve the infrastructure (including IT security) that make these initiatives possible. This is our constant struggle."

~ Manager, IT  
North American food service  
company

their sales personnel were able to identify high value prospects, quickly close deals and meet their yearly sales quota. Companies without trust in their data reported only 59% of their team meeting this target.

**Figure 4: High Trust and Customer Interaction**



Source: Aberdeen Group, December 2012

Customers purchasing from organizations with trustworthy data were also more likely to be satisfied with the experience and willing to refer them to others, which is reflected in a higher net promoter score. Organizations that used trusted data in their sales cycle reported a 74% NPS score, while companies without reliable customer data reported only 68%.

### Fewer Data Errors

A key component to building and maintaining trust in business data is keeping information largely error-free, and quickly fixing problems when they do occur. Companies without the dedication or tools to provide trustworthy data are shown to have significantly more issues with their data, as well as needing to spend considerably more time and resources correcting them.

For instance, the average organization in the [Master Data Management](#) research study reported having over 6.2 million records across all their data domains, including customer, product, material, supplier, accounting, and human resources. As Table 4 shows, organizations with information governance tools and high levels of trust had over half a million fewer records with significant errors.

#### Definitions:

√ **Net promoter score (NPS)** is calculated by taking all satisfied customers willing to refer a company and subtracting customers that are unsatisfied and unwilling to give a referral. For instance, if 75% of an organization's customers would give a referral, and 25% would not, their NPS would be 50%.

**Table 4: Fewer Errors, Faster Correction**

Category	Number of records	Percent of records with errors	Number of incorrect records	Hours needed to fix one error	Total hours needed to fix all errors
Average company with high levels of trust in data	6.2 million	11.8%	732,000	5.8 hours	4,246,000
Average company without high levels of trust in data		20.9%	1,296,000	7.4 hours	9,590,000
<b>Performance benefit for companies that trust their data</b>			<b>564,000 fewer bad records</b>	<b>1.6 fewer hours to fix one error</b>	<b>5,344,000 fewer hours</b>

Source: Aberdeen Group, December 2012

High trust also correlated with faster error remediation, which as seen above, meant that organizations with this high level of trust would spend an average of 5 million fewer man hours to fully cleanse their data environment. To put that into comparison, a single employee at a low trust company working to improve their data up to the level of a high trust company could have started when Julius Caesar ruled Rome in 45 B.C., worked 40 hours a week year after year until the current date in 2012, and they would still have a more than 600 years left.

Obviously, no organization would tackle a data-cleansing project of this magnitude using manual means, and no company needs perfect, completely error-free data. However, in order to get data to a reasonable level of reliability, companies with such a severe problem would have to employ a third party company to provide data quality services, or purchase automated data quality solutions to complete the task in large batches. The time and expense involved with either of those options are not inconsiderable, and could be largely avoided by addressing the problems at a smaller scale over time. The power of information governance tools is not that they will prevent all possible errors, but rather that they make even massive amounts of data manageable and ensure the most important data elements can be relied on.

"The Board of Directors have to be persuaded to treat data as an asset, from acquisition, safeguarding, ownership, through disposal. When there is a disconnect between top management and the users of data (particularly unstructured), the data strategy becomes unfocused..."

~ IT Director

Mid-sized IT Services Company

Asia/Pacific

## Summary and Key Takeaways

Making sure business data is accurate and reliable has long been a direct contributor to good organizational performance. However, with the age of Big Data upon us, and information growing at 56% year-over-year, data governance has changed from a long term nice-to-have to an immediate must-have. Ignoring issues of data security, data quality, and data integration costs organizations millions, hurting their efficiency and losing them valuable customers. These headaches are only going to get worse down the road and require more costly solutions. Companies desperately trying to fix their poor decision making and out-of-control data costs, as well as companies

looking to embrace Big Data and establish a firm data-driven competitive advantage, can find the answer in information governance solutions. In particular, Aberdeen recommends organizations consider the following steps:

- **Evaluate data needs and available budget.** The current economic downturn is a major reason why many companies have not invested in information governance in recent years, so it is important to consider the financial realities when finally taking the plunge. An audit of existing data sources and end-user requirements will help identify what elements of a governance strategy should be addressed first. A frank examination of the budget will identify what you can afford. However, remember that improving data governance can benefit the line-of-business as well as IT. In many cases, budget from departments like marketing and sales is being used to fund data initiatives.
- **Get in the game.** Regardless of financial flexibility, organizations struggling with data quality must take steps to improve their situation as soon as possible. Consider implementing small projects for a single, essential data domain or business unit, then use that success story to roll out to other areas.
- **Don't neglect data security.** Of the many elements of data governance discussed in this report, the one with the greatest potential to seriously harm a company is a major data breach or exposure event. Even the best organizations experience data loss and audit penalties, and all it takes is a single headline to destroy a reputation and lose customers. Protecting sensitive data requires diligence, a strong security solution, and constant investment to keep current with the latest threats and regulations.

For more information on this or other research topics, please visit [www.aberdeen.com](http://www.aberdeen.com).

### Related Research

[\*Data Management for BI\*](#); January 2013  
[\*Big Data for Small Budgets\*](#); December 2012  
[\*Agile Data Integration: The SMB Perspective\*](#); December 2012  
[\*Maximizing the Value of Analytics and Big Data\*](#); September 2012  
[\*The State of Big Data: Video Benchmark\*](#); August 2012  
[\*Agile or Fragile? Your Analytics, Your Choice\*](#); July 2012  
[\*Beyond Agile Analytics: Is Agile Data Integration Next\*](#); June 2012  
[\*The Ultimate Storage Tier\*](#); May 2012  
[\*Managing the TCO of BI: The Path to ROI is Paved with Adoption\*](#); May 2012  
[\*Enabling Access to Big Data with Data Integration\*](#); April 2012  
[\*High Performance Organizations Empower Employees with Real-Time Mobile Analytics\*](#); April 2012

[\*Mobile BI 2012: Accelerating Business on the Move\*](#); March 2012  
[\*The Little Elephant in the Big Data World: Hadoop 1.0 Goes Live\*](#); March 2012  
[\*Divide & Conquer: Using Predictive Analytics to Segment, Target and Optimize Marketing\*](#); February 2012  
[\*Operational Intelligence - Part 1: Driving Performance with Tactical Visibility\*](#); February 2012  
[\*In-memory Computing: Lifting the Burden of Big Data\*](#); January 2012  
[\*The Role of Big Data Analytics in HR: Speed, Satisfaction and Scale\*](#); January 2012  
[\*Agile BI: Complementing Traditional BI to Address the Shrinking Decision-Window\*](#); November 2011  
[\*Agile BI: Three Steps to Analytic Heaven\*](#); March 2011

Author: Nathaniel Rowe, Research Analyst, Enterprise Data Management ([nathaniel.rowe@aberdeen.com](mailto:nathaniel.rowe@aberdeen.com))

For more than two decades, Aberdeen's research has been helping corporations worldwide become Best-in-Class. Having benchmarked the performance of more than 644,000 companies, Aberdeen is uniquely positioned to provide organizations with the facts that matter — the facts that enable companies to get ahead and drive results. That's why our research is relied on by more than 2.5 million readers in over 40 countries, 90% of the Fortune 1,000, and 93% of the Technology 500.

As a Harte-Hanks Company, Aberdeen's research provides insight and analysis to the Harte-Hanks community of local, regional, national and international marketing executives. Combined, we help our customers leverage the power of insight to deliver innovative multichannel marketing programs that drive business-changing results. For additional information, visit Aberdeen <http://www.aberdeen.com> or call (617) 854-5200, or to learn more about Harte-Hanks, call (800) 456-9748 or go to <http://www.harte-hanks.com>.

This document is the result of primary research performed by Aberdeen Group. Aberdeen Group's methodologies provide for objective fact-based research and represent the best analysis available at the time of publication. Unless otherwise noted, the entire contents of this publication are copyrighted by Aberdeen Group, Inc. and may not be reproduced, distributed, archived, or transmitted in any form or by any means without prior written consent by Aberdeen Group, Inc. (2012a)