White paper

Create your ideal data quality strategy

Become a more profitable, informed company with better data insight

Table of contents:

Introduction

Organizations are using data more than ever before. This valuable business asset is a competitive differentiator being used to influence an increasing number of business decisions, improve operations, better understand consumers and so much more.

To use data in this way requires an accurate, consolidated and complete data source for all customer information. Without that level of access, organizations simply are not able to gain the level of insight needed at a quick enough pace.

The right data management approach is needed to achieve this centralized and accurate resource. Unfortunately, many organizations today are falling behind on their strategy. Today, 78 percent of U.S. companies could improve their level of data management sophistication.

New research from Experian Data Quality shows that companies with a more sophisticated approach to data management see a lesser degree of inaccurate information and experience higher profits as an organization. If organizations can find the right data strategy that works for them, they can achieve greater data success.

The level of sophistication has to do with the people, processes and technology surrounding data. Any one of those areas can be a short fall or strong point that can dramatically affect the quality of data and how much an organization is able to use it.

This resource looks at the various levels of data quality sophistication, current challenges with data management strategies and reviews how changes to the people, processes and technology surrounding data management can have a dramatic effect on an organization's ability to gather insight from data.

The levels of data quality sophistication

Data quality and data management sophistication varies greatly depending on the organization. There are a variety of factors that go into the level of sophistication, but we see four key segments of sophistication. Each of these takes into account the people, processes and technology surrounding data management.

As each of these progress from unaware to optimized, there are key changes in people, processes and technology. For people, there is increasing ownership and centralization around a common goal. For processes, data quality becomes more of a standard part of business practices, rather than looked at on an ad-hoc basis. Around technology, centralized and sophisticated data quality tools are implemented and there is a move away from manual processes.

To become sophisticated in their approach, organizations need to review data management holistically. Many organizations analyze data based on individual departments or localized regions. That view promotes one-off fixes and data standardization for just that individual department's needs. For organizations to be really data-driven and make effective decisions based on that data, they need to review data and data management from a more holistic perspective.

The level of sophistication varies pretty dramatically across businesses today and the bulk of organizations are at different stages in each of these key areas.

Unaware

- Understanding of data quality impact is patchy around the business
- Data quality fixes sometimes happen
- Excel or manual processes are the primary data management methods used around the organization

Reactive

- Good knowledge of data quality impact, but no data-specific roles exist within the business
- Data quality fixes happen, but in departmental silos
- Excel or manual processes are the main data management methods, but departments have sophisticated tools

Proactive

- Data quality sponsors exist and success metrics are outlined
- There is clear data quality process ownership between business and IT
- Technical aspects focus on discovery and root cause analysis

Optimized

- There is a central data role, such as a CDO, in place that is accountable for corporate-wide data assets
- Data quality is monitored as part of standard business practices
- There is a platform approach to profiling, monitoring and visualizing data

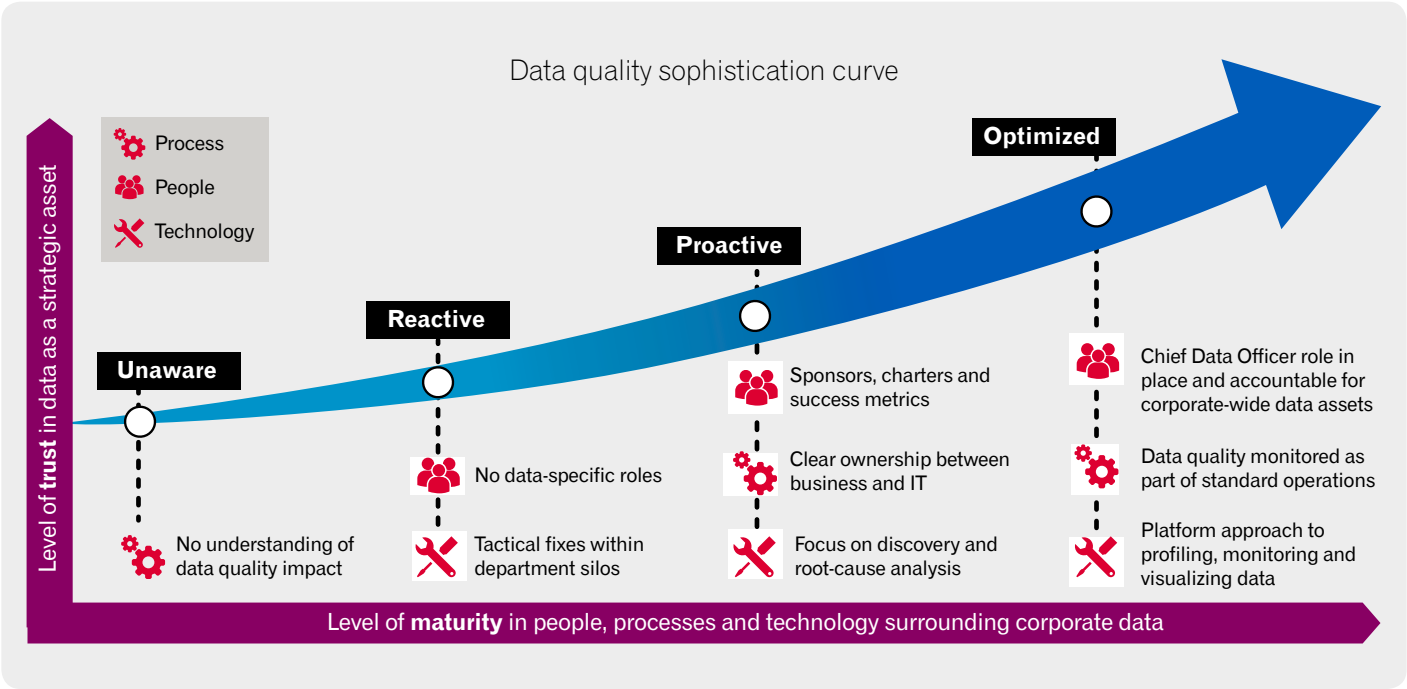
Data quality sophistication curve

Today, most organizations are at lower levels of sophistication when it comes to their data management strategy. From a recent Experian Data Quality study, we can see that while the bulk of organizations are investing in data quality, most still do so in a siloed fashion. That means the central approach needed for more sophisticated data management strategies is not in place.

Just one in four organizations today have a highly sophisticated approach to data quality, falling in the optimized category. Twenty-nine percent are proactive, 29 are percent reactive and 17 percent are unaware.

Interestingly, larger companies tend to select more sophisticated options. That could be due to their ability to invest in larger data quality solutions or simply that the amount of data they host has become so large that they have been forced to maintain it centrally.

This shows how much companies vary in terms of their strategy and how they approach data management today. While we see fewer companies in the unaware phase than any other category, there is still a great deal of room for improvement.



Improving data management strategy

To move up in terms of data quality sophistication, organizations need to take a methodical approach to improving their data management techniques. That revolves around the three key areas of people, processes and technology.

All three of these areas need to be advanced to achieve the right level of data accuracy, consolidation and accessibility required to stay competitive in the data-driven environment. The subsequent sections walk through each of these areas and the material improvements that can be made to advance data management maturity.

People

People are the backbone of any data management strategy. They develop the strategy, implement it and ensure it is adopted by the organization. Without the proper people, data management strategies cannot move forward. There are many issues in today's structure that can be learned from and improved to advance data strategies.

Issues in today's structure

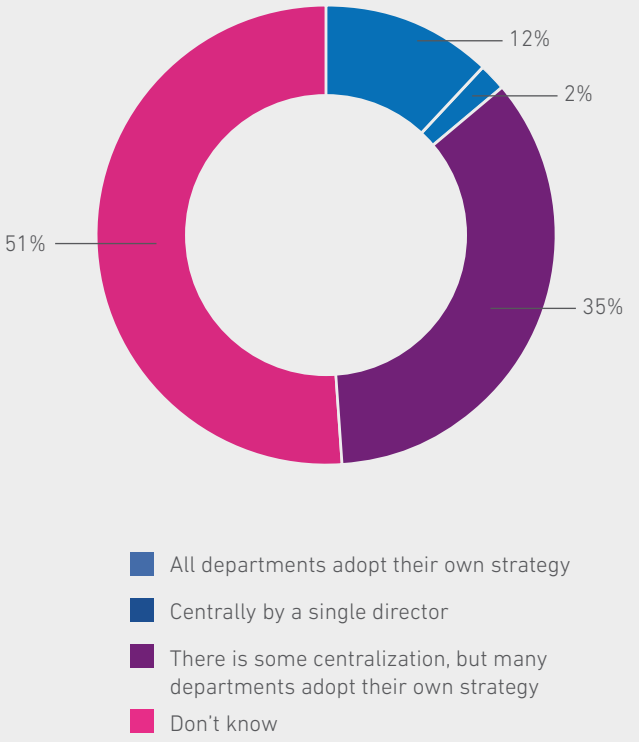
Among those with a data quality strategy in place, very few manage that strategy centrally. Just 35 percent say data quality is reviewed and maintained centrally by a single director. To be optimized in their data quality strategy, companies need a central owner and even for a proactive level strategy, a high degree of centralization is needed. Most commonly, companies report that there is some centralization, but that many departments still adopt their own data quality strategy

Interestingly, there is no systematic difference between to company size or the number of individual databases a company maintains and how they review and maintain their data quality strategy. That means everyone is in about the same place when it comes to their data quality structure.

In the past two years of reviewing this data, only a modest improvement has been made in the area of centralization. There has been a slight increase in the proportion of companies that manage their data quality strategy in a centralized way with ownership resting with a single director, up from 30 percent in 2014 to 35 percent in 2015. Organizations still have a long way to go in this area.

Organizational structure around data is so important that it even correlates with company profits. More companies who have enjoyed a significant increase in profits in the last 12 months manage their data quality strategy in a centralized way, with ownership resting with a single director.

Chart 1
Centralization of data management



Hiring the right people

To advance a data management strategy, organizations need to centralize ownership. That means transferring departmental, disconnected strategies into one central strategy that works for the entire organization.

Among those with centralization, ownership most commonly rests with a chief data officer or the CIO/CTO. While the chief data officer is a relatively new title in the past three years, there is certainly a case for adding a CDO to the organization, especially considering the value of data and the benefit of having someone to take responsibility for the quality, standards, meaning, security, metrics, integration and coordination of data among the various divisions

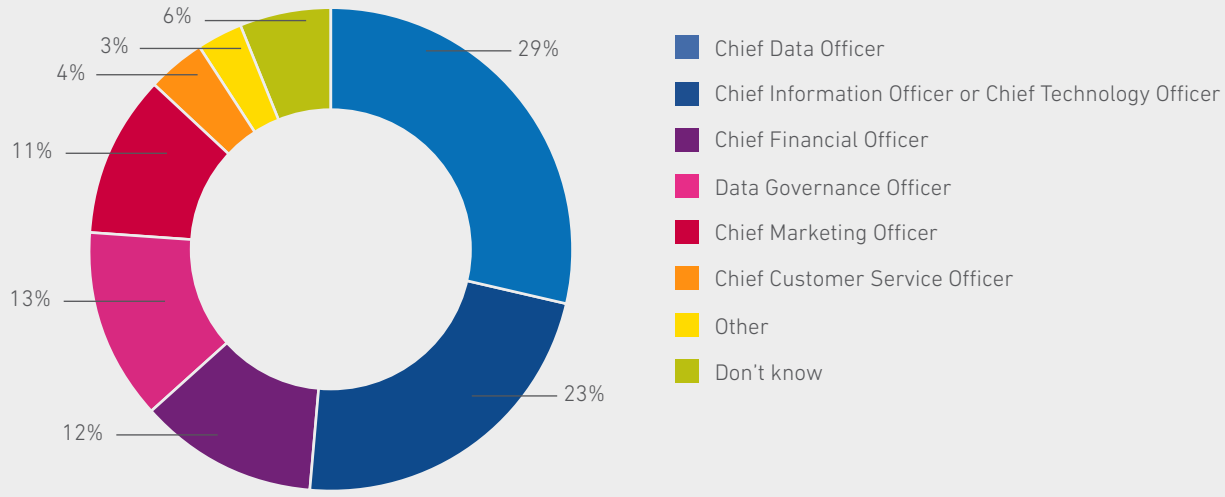
While having a central director is of vital importance, so are additional members of a data team that need to be assembled. Data professionals such as: business data stewards, data service officers, analytics professionals, data scientists, records managers, etc., all need to help the central owner enforce policies and promote the use of data across the business.

Having a team dedicated to ensuring the accuracy, security, standards, metrics and coordination of data across the business will have a big impact on the usability of information and the actionable insights organizations are able to derive.

Did you know?

Organizational structure around data is so important that it even correlates with company profits. More companies who have enjoyed a significant increase in profits in the last 12 months manage their data quality strategy in a centralized way, with ownership resting with a single director.

Chart 2
Owner of central data quality strategy



Processes

Processes are essential to ensuring data management functions the way that it should within the organization. They need to be consistent to ensure information is treated the same way, regardless of channel or department, and proactive to ensure information is ready to be used whenever the business requires.

Reactive processes inhibit insight

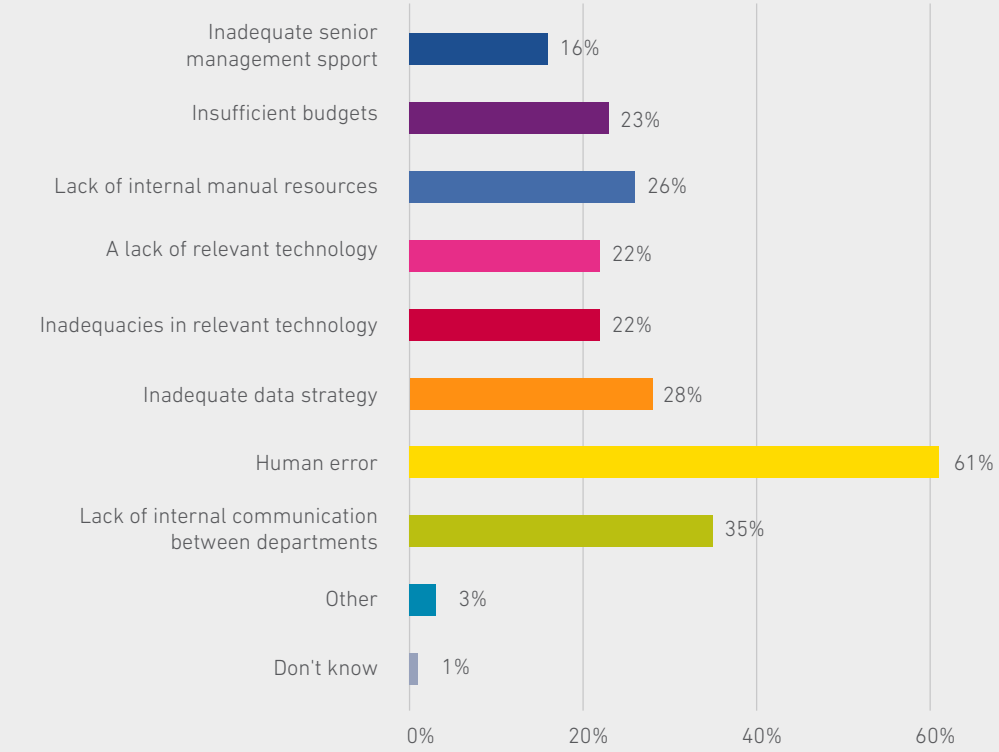
The processes surrounding data management are very segmented and reactive. We can see this in several ways from the data.

First, it's a human problem: research shows that among those with contact data accuracy issues, the most common cause is believed to be human error. This explains why companies and departments of all different sizes are affected. No matter what the department, organizations will have a person enter data at some point within their processes. Without a central technology or data governance process in place to prevent this, human error will continue to cause problems for organizations. Because of the segmented strategies, today, very little is being done to prevent this type of error, which is why it affects 61 percent of organizations.

Did you know?

Today, 93 percent of companies state that they are making an effort to discover and find data quality issues across their company. However, the majority, 57 percent, say data quality issues are detected when they are reported by employees, customers or prospects.

Chart 3
Reason for inaccuracy



Second, disparate processes appear when looking at monitoring and discovering data quality issues. Today, 93 percent of companies state that they are making an effort to discover and find data quality issues across their company. However, the majority, 57 percent, say data quality issues are detected when they are reported by employees, customers or prospects. That means fewer than one in two companies conducts proactive data audits to discover data quality issues. These processes again come with data quality sophistication and central ownership.

Finally, we can see a lack of processes when we consider what organizations find challenging about managing data quality. Ninety-two percent of companies say they find some element around managing data challenging. The most significant seems to be fixing data quality issues before they negatively impact the business. Again, not having a proactive data strategy means the right processes are not in place.

Switching to central, proactive processes

There are a number of advanced processes organizations can put in place to further their data management strategy. A big part of moving a strategy forward is having central processes that are consistent across the organization and are part of daily activity.

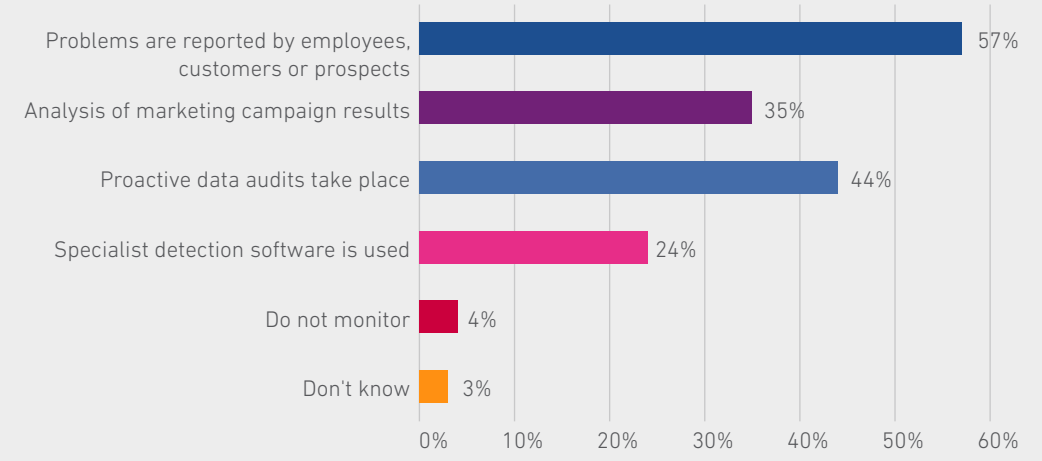
There are several key central components that need to be considered. First,organizations should think about what information they want to collect and what is valuable.

Attempting to collect data for data's sake is never a good strategy. The volume of information is increasing at such a rapid rate organizations simply can't keep up with making sure all structured and unstructured data is ready to use at any moment. Organizations should focus on the primary data sets used across the company.

Second, take a proactive approach. The data shows that over half of companies take a reactive approach to data management in some way or another. It could be with their entire data strategy or in fixing data quality issues. Either way, reactive is not a good place to be when it comes to data quality. If issues are only found after negatively impacting the business, there could potentially be many uncovered errors causing further harm before they are discovered. Look at more preventative processes and measures to ensure information is accurate when it is used.

Finally, it is important to keep in mind that the technology and business side of the organization need to work together. While technology certainly plays a key role in data management, data needs to be viewed from a business lens. The primary data stakeholder needs to have a firm grasp of the purpose of data within the organization and how the business plans to use it. Then they can come up with management practices that fit that need. The disconnect between technical and business users is another primary reason for the high degree of inaccurate information.

Chart 4
Discovery of data quality errors



Technology

To eliminate human error and reduce the high degree of inaccuracy that exists today, organizations have to use technology. The volume of information is too high today to rely solely on manual processes. Therefore, when technology is implemented, organizations need to be sure the investment they make has the biggest impact on the organization.

Divided technology

The bulk of organizations are using technology today for data quality and most plan to continue their investment in data quality technology over the course of the next 12 months. In fact, 88 percent of companies have some sort of data quality solution in place today.

However, the types of tools in place vary greatly depending on the organization. Most of the major types of data quality

Chart 5
Used data quality tools

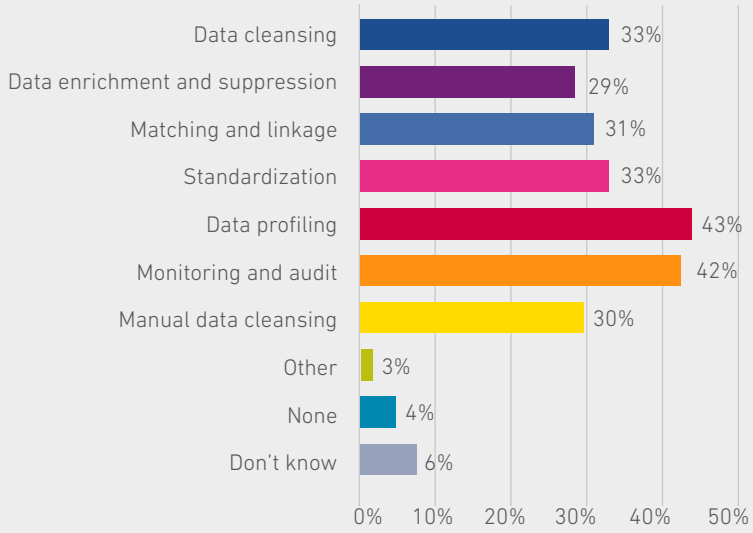
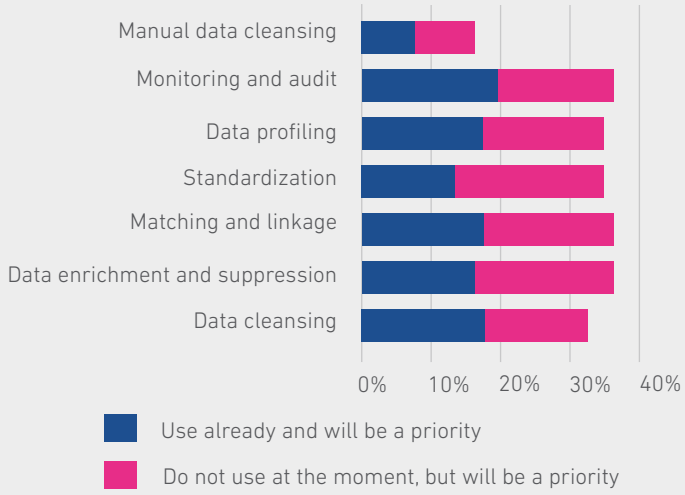


Chart 6
Data quality tool investment



tools are used by a third of companies or fewer, suggesting organizations take a very varied approach to the types of solutions they have in place.

The most widely used data quality solution is monitoring and audit technology. This is followed closely by data profiling technology and matching and linkage technology. Twenty-nine percent use data cleansing technology, but the same percentage use manual data cleansing, a technique indicative to lower levels of data management strategy.

In addition to a varied approach, the research also shows that the tools, much like the organizational structure, are departmentalized. While there are some fairly sophisticated techniques being used, the variety of techniques shows that not only differences exist between organizations, but also between internal departments.

Did you know?

The most widely used data quality solution is monitoring and audit technology. This is followed closely by data profiling technology and matching and linkage technology. Twenty-nine percent use data cleansing technology, but the same percentage use manual data cleansing, a technique indicative to lower levels of data management strategy.

Preventative technology consistently installed

Data quality technology is a cornerstone of any data management strategy. Given the big data environment, organizations need technology to deal with the volume and speed at which information is coming in.

Moving through the levels of data quality sophistication, organizations need to step away from purely manual or departmental technology and attempt to implement solutions consistently across the organization. If one department is using an address verification solution, another department should not be using a different solution to do the same thing from a competing vendor or not use a solution at all.

Information needs to be consistently maintained, standardized and validated across the organization. Tools should be consistently selected and implemented to make sure data is treated under the same technology.

That centralization also tends to lend itself to more advanced techniques. When resources are pooled together, it is more likely that basic technology can be implemented more cheaply than with separate purchases across the organization. This means more resources can be freed up to look at data monitoring or even visualization.

Interestingly, the data shows a correlation between the sophistication of a company's data quality strategy and the number of tools they use. Companies who have implemented a more sophisticated approach to data management, such as optimized or proactive, do in fact use a wider variety of data quality tools.

Finally, technology should be implemented in a preventative fashion. It used to be that many companies executed quarterly audits to manage their information. In today's environment, that is no longer enough. Look at more proactive technologies that can validate and standardize information as it is entered and then monitor it over time to ensure continued usability.

Conclusion

As organizations look to better use the valuable asset of data, it is essential that they find the right data management strategy to ensure it provides value for the organization. Data can be used for good in many ways, but if not properly maintained, it can cause damage.

Most organizations are at lower levels of data quality sophistication at this stage. But as investment continues and the CDO continues to become more popular, organizations will inevitably advance their strategies into more central functions. The people, processes and technology around data need to operate in a more coordinated fashion to ensure consistency and usability across the business.

While this type of strategy cannot be achieved over night and there is not magical piece of software that can solve all data quality problems, organizations can take small steps towards improving their data management. It just takes time and coordination to move the needle in the right direction.

Methodology

In December 2014, Experian Data Quality commissioned a research study to look at current approaches to data quality. This data quality benchmark reviews the evolution of data quality and consumer interaction while providing best practices for data management.

Over 1,200 respondents globally took part in the research, produced by Dynamic Markets for Experian Data Quality. Individuals from the U.S., UK, France, Germany, Spain and Australia completed the survey. Industry sectors included in the sample were finance, public sector, retail, manufacturing, utilities and education. Respondents consisted of C-level executives, vice presidents, directors, managers and administrative staff connected to data management, across a variety of functions.

Experian
53 State Street
Boston, MA 02109
T: 1 888 727 8822
www.edq.com