Best Practices for Monitoring Data Quality

Improve Database Effectiveness with Accurate Data

An Experian QAS white paper

March 2013
State and local government agencies face a multitude of challenges in today’s environment and data accuracy is more important than ever. As the speed of business is constantly increasing, decisions are made rapidly and impacting the ever increasing channels of communication.

Data accuracy is a major area of focus for most organizations. Despite the overall advances in analytics and business intelligence, most agencies struggle with data accuracy. According to a 2013 Experian QAS survey, 94 percent of organizations believe there is some level of inaccuracy within their systems.

To prevent inaccuracies and improve database performance and reporting, it’s more important than ever for agencies to monitor data quality.

Measure the effects of inaccurate data
Inaccurate contact data requires organizations to face the consequences of poor data quality. Inaccurate data makes it harder for agencies to adhere to data governance policies. In addition, the Experian QAS survey found that 91 percent of organizations think that at least some of their departmental budget was wasted in the past 12 months as a result of contact data inaccuracies. On average, 12 percent of departmental budget was wasted.

It is worth noting the correlation between number of distinct databases within an organization and amount of budget thought to be wasted – more databases indicated more wasted budget.

There are other consequences facing companies. 93 percent of organizations say they have been negatively impacted in some way over the past three years as a result of contact data accuracy issues. The most common problem reported is sending mailings to the wrong address. This is followed by staff inefficiencies. 32 percent of respondents said that customer perception is negatively influenced by inaccurate contact data. Additionally, 29 percent stated that they had lost a customer because of inaccurate data input. All of these problems damage the bottom line, analytics and ultimately hurt the customer experience.

Gain corporate stakeholders
To monitor and measure data quality, it is important to identify data quality stakeholders within the organization. Identifying data quality stakeholders among different departments can help to make incremental improvements in data quality across the organization while prioritizing projects.

As the data quality stakeholders determine and make a business case for data quality improvements, there are several considerations for ensuring success and executive buy-in. First, it’s important to make the proposal credible. Stakeholders need to demonstrate the data quality project will provide tangible benefits to the organization. Be sure a proposal includes financial models with a return on investment.

Next, it’s important to demonstrate additional benefits. In addition to the bottom line, many senior managers look for added incremental benefits to the organization’s efficiency. By tying data quality improvements to strategic initiatives helps to make a compelling business case. Understand if there are cost savings plans, compelling events or other initiatives that data quality can positively impact. It’s important to leverage other stakeholders inside the organization when creating a data quality proposal.

Prevent human error
The main cause of data quality problems is human error, which was cited by 65 percent of organizations in the Experian QAS study. To operate effectively, organizations...
need to do more than just exist in each channel; they must create a seamless customer experience across all channels.

To conduct business effectively across channels, organizations need accurate data for analytics. Business intelligence is only as accurate as the information that supports it, and as mentioned previously, managing that information is challenging for many businesses.

In order to improve data accuracy, businesses need to eliminate human error, the main cause of poor data quality. There are several steps businesses can take to remedy this issue.

**Identify data entry points**
Businesses need to understand how information enters their system and through what means. Consider all channels and data entry points so a full data workflow can be created. Then, prioritize projects based on high volume channels and where excessive data quality errors occur.

**Train staff**
Staff education can go a long way toward improving data quality, considering a lot of information is still manually entered by employees. Explain the importance of accurate data to employees and educate them about how information is used throughout the business. Implement quality control checks and accountability for data accuracy for employees.

**Utilize automated tools**
Next, organizations should utilize automated tools to verify data and ensure accuracy. Automated processes reduce staff bandwidth requirements. Software solutions can be implemented across various channels to help prevent the capture of inaccurate contact information and standardize important data elements like the address. Identify what data elements are most important to operations, then evaluate and prioritize available solutions.

**Continuously clean data**
Finally, incorporate technology that continues to clean
information periodically. Even with software tools working at the point of capture, consistent database maintenance is essential. Regular cleansing allows organizations to review information and ensure that tools are still effective in managing the data to the expected level of quality.

**Reduce duplicate data**
Duplicate data has become one of the most common data quality errors for organizations. The Experian QAS study found 92 percent of organizations admit to having duplicate data. Duplicate information spreads account history across multiple records. This impedes intelligent decision making and can harm the customer experience.

Duplicate records are created in a number of different ways. The majority of survey respondents blamed human error and multiple points of entry. Other common responses include issues with multiple databases and multiple business channels.

Whatever the cause, it is important that organizations remove duplicates from their database in order to achieve efficiency and business intelligence goals. There are several techniques organizations can use to remove existing duplicate records.

First, organizations should standardize contact data. Since contact information is typically found in every record, it can be used to help identify duplicate contacts.

Next, administrators should define matching criteria, as well as the tolerance level for what is considered a duplicate record. It is important to have an outline of what a single record means for the organization before the process of merging records begins.

Software should then be used to identify duplicates based on the defined criteria. While manual review is preferred by some organizations, it is important for larger organizations to utilize software to ensure duplicates are identified according to the given definitions. Once records are identified, then the primary record can be determined and the merge purge process can begin.

Once duplicates have been removed, it is important that organizations put processes in place to reduce the possibility of additional duplicate records. Fuzzy matching technology can help reduce that trend.

Fuzzy matching technology uses computer-assisted translation to link records that may be less than one hundred percent exact. Most CRM systems require an exact match to find an existing record, while fuzzy matching allows systems to identify similar records. By utilizing this software, fewer duplicate records are created.

**Conclusion**
There are steps agencies can take to improve the data quality monitoring process. Accurate analytics will allow businesses to make more informed business decisions, operate more efficiently and improve customer satisfaction.

Monitoring data quality is an integral step for improving CRM ROI. Stakeholders should ensure the strategy they have in place for data quality is producing the required results.