



**The Power  
of  
Zero Based  
Data Solutions™**

## Freedom from Data Quality Issues with a Comprehensive Zero Based Open Source Solution.



### The Missing Piece

Data is the fundamental building block of every business, data in the form of client information, sales information, contact information etc. fuels the operation of every business. In today's business environment which enables data entry from multiple points and through a myriad of processes, data quality has become an increasing concern for businesses trying to succeed in an ever more competitive environment. If your business has Data Quality issues you are not alone, The Data Warehousing Institute reported that 50% of its respondents felt that actual company data quality is worse than the organization overall thinks. Furthermore, more than half of respondents indicated their organizations had suffered losses due to poor data quality.

Data fuels your organization, poor data quality can restrict the implementation of mission critical initiatives such as ERPs, CRMs and SFAs. Additionally, poor data quality obscures the view of an organization through clouding critical client information, employee information and financial information. Data quality problems can lead to missing additional revenue opportunities, risking regulatory issues and forfeiting the intelligence gained from a clear view of the data in our business. Data quality is the core of your business operations, focus on your data and focus on your company's future.



### The Critical Piece: OpenDQ

#### Data Quality Issues: Addressing the sources and their implications

The prevalence of today's flexible data entry options from cell phones to web sites and expanding mobile devices are increasing the opportunity of data quality issues to be introduced in enterprise data. Data quality issues can stem from many sources including:

- Poor Data Entry
- Missing Data Fields
- Multiple Databases with disparate database rules
- Legacy and other older systems with obsolete or unknown data
- Merging of data from acquisition or other business units with separate business rules

Understanding and addressing the sources of data entering organizations and creating standards around data entry is key to preventing future data hygiene issues. This is best accomplished through a thorough analysis and mapping of all points of entry and a systematic plan for addressing current issues, while implementing processes to address future data quality issues.

The first step in seeing the road ahead clearly is acknowledging the problem and implementing a solution. In the past solutions to address data quality were complex, lengthy and expensive often involving hardware, software licenses and term commitments. With Infosolve Technologies Zero based Data solutions, data quality for your organization can be achieved with speed, accuracy and cost efficiency. Infosolve Technologies OpenDQ solution is an open source solution that provides organizations with a powerful data quality solution without the need for upfront hardware, software or term contracts.

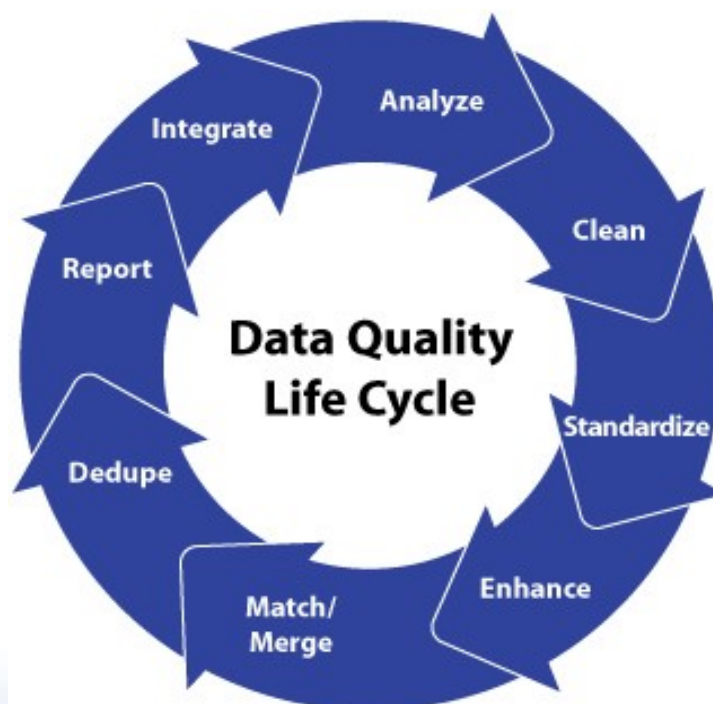
### **The OpenDQ Solution: Focusing on the complete data quality life cycle.**

#### **Comprehensive. Easy to Use. Cost Effective.**

The OpenDQ solution is a thorough integrated data solution that provides data quality coverage for your entire enterprise. OpenDQ can assist your organization in identifying quality problems and taking corrective actions to address these issues now and in the future.

With OpenDQ your organization will have the power to develop, deploy and monitor enterprise wide data quality programs, while preventing new problems from occurring. The OpenDQ solution integrates all of the key components of data quality with ETL (extract, transform and load) software creating unmatched efficiency in process control, data flow, and data processing.

### **OpenDQ manages the data quality life cycle of your data through a comprehensive process.**



## **Data Analysis**

In the data analysis phase, your data is analyzed by the OpenDQ data profiler to identify the issues residing in the data. This process typically identifies critical issues including:

- Hidden business rules
- Missing or incorrect column values
- Pattern and format issues
- Data value distribution
- Minimum, maximum and average values
- Missing records and orphan records
- Master-Detail relationships
- Duplicate records

At the close of this process, a clear picture of the data quality state of the data set and the issues that need to be addressed in the data quality solution are identified.

## **Data Cleansing**

After the analysis is complete, the cleansing of the data can begin. OpenDQ's powerful data cleansing components addresses complex and difficult data issues with comprehensive data cleansing tools including:

- Noise elimination
- Spell check and correction options for text fields
- User updated dictionaries for standardization
- Address verification/correction interface for the U.S. and 240 other countries (Subscription needed)

## **Data Standardization**

OpenDQ offers complete standardization of all types of data based on businesses specific rules, including:

- Name, Address, Inventory, Vendor, Accounting and Product Data

## **Data Enhancements**

In the Data Enhancement stage OpenDQ, enhances and organizes existing data by utilizing internal or external reference sources. OpenDQ can support:

- Database and file based field lookups
- Custom dictionary lookups with fuzzy match options

## **Data Matching**

OpenDQ can handle your most complex record linkage issues including single source matching and matching data across multiple sources. Matching can be performed using exact or the latest approximate string matching techniques such as hybrid algorithms, heuristic algorithms and edit distance algorithms.

Open DQ Match processes include:

- Matching across multiple columns
- Matching across multiple data sources
- Master data set matches (especially important with incremental comparisons)
- Flexible match tolerances for columns or for the entire record.

## **Data De-duplication**

In the De-duplication phase, duplicate data records are removed. For exact duplicates where the records are identical, de-duplication is a simple process, however most data contains inexact duplicate data that requires approximate string or fuzzy match logic to identify the duplicates in the data.

Once duplicate entries are identified the records can be merged or removed by OpenDQ based on specified business rules.

## **Data Reporting**

Throughout the entire Data quality life cycle OpenDQ provides comprehensive data profiling information that chronicles the quality and accuracy of the data.

## **Data Integration**

OpenDQ has simplified the process of Data Integration through the integration of an enterprise grade ETL tool with the data quality components. With this integrated ETL tool, data can undergo extensive extraction, transformation and integration within a comprehensive data quality management framework.



## The Complete Solution: Puzzle Solved

### The power of Zero Based Data solutions driving your business.

Open DQ is built for simplicity and ease of use on an open source platform that enables users to have access to management features with point and click functionality. The OpenDQ solution includes functionality that is easy to use and highly flexible including:

- Comprehensive Reporting to monitor data quality
- Custom rule development to flag non conforming records
- Flexible processing methods including batch or interactive
- Work flow and scheduling capabilities
- Integration options into other enterprise data processes
- Drag and Drop components

The clean and accurate data you need to fuel your business now and in the future is achievable and accessible with Infosolve Technologies Zero Based Data solutions -- with no upfront hardware, software, or term commitments plus the Zero Defect Data Guarantee™, assuring no new defects will be introduced into your data. Additionally, OpenDQ features the rapid, expert implementation of Infosolve Technologies experienced data quality teams, assuring that your needs are addressed with speed and your ROI is achieved. OpenDQ is the only data quality solution your organization will need to meet its data quality needs now and in the future.

**Experience the power of immediate, accurate data today with OpenDQ!**