



***Q3 technologies***

## Going an extra mile to create affordable, cutting edge components of future technology

Q3's global sourcing model gives the maximum benefit to customers in terms of cost savings, improved quality, access to highly talented professionals, flexibility of operations and reduced time to market.

## Case Study - Client in Chicago, IL

### Technical Situation

Client is the industry's choice for Companies that often cannot rely on the information that serves as the very foundation of their primary business applications. Inaccurate or inconsistent data can hinder the company's ability to understand its current – and future – business problems. This leads to poor decisions that can cause a host of negative results, including lost profits, operational delays, customer dissatisfaction and much more.

Compliance expertise, acquired during more than two years of operating success, positions client to help provide the infrastructure to transform raw data into consistent, accurate and reliable corporate information.

Client delivers software solutions with emphasis on Data Profiling, Data Cleansing, Data Quality and Customer Data Integration. A Chicago-based privately held Company in its 8th year of business.

Client provides industry-leading technologies that allow organizations to collect profile and cleanse their corporate data. These are essential processes for corporations that are involved with major projects such as customer relationship management (CRM), enterprise resource planning (ERP), business intelligence systems and data warehousing.

Client is Microsoft partner in providing data profiling and data cleansing systems which are seamlessly integrated with Microsoft SQL Server Integration Services on IBM DB2 and Microsoft SQL Server repositories in Context of taking data management to the next level.

### Client Profile

### Product Description

The product can access most databases and file formats and profile the results accurately as well as to greatly reduce the time and effort to move and integrate data.

Also, the software is an effective data quality strategy that can help you better understand your business environment, allowing you to maximize profitability and reduce costly operational inefficiencies.

Our technology platform is primed for delivering a wide array of data services that offer maximum flexibility, exceptional quality and execution speed. Our technology platform combines quality processes, cutting edge best practices, specialized hardware, data integration tools and a vast repertoire of reusable code libraries into a unified framework that delivers a highly optimized solution to our clients.

A very critical part of our technology is our quality processes and methodologies. Every process step in our implementation is subject to our stringent internal quality checks. These quality checks result in a solution that maximizes quality and reduces errors.

- » Quality processes and procedures that minimize errors.
- » GUI based development, maximum flexibility, rapid execution and accuracy.
- » Specialized hardware, capable of handling data loads of any size ranging from a few megabytes to terabytes of data.
- » Reusable code libraries and templates, support rapid roll out.
- » Web based project management/bug tracking software, enhances Collaborative effort and communication.

## Overview

Poor data quality can play a significant role in eroding your company's bottom line and competitive advantage in the marketplace. Data corruption can occur due to a number of reasons; the following are the most common reasons that contribute significantly to poor data quality.

- » Infusion of non conforming data during data migrations, data conversions and data acquisitions.
- » Flawed and inconsistent data validation at the point of entry.
- » Non standard data modification e.g. changing data on databases directly.
- » Erroneous data entered by users in free form data fields such as web forms, user interface screens etc.

## Business Situation

The fair lending process – Automation was required to help perform complex statistical regression. The client wanted us to introduce five constructs to accommodate all changes in data format and content. They are as follows:

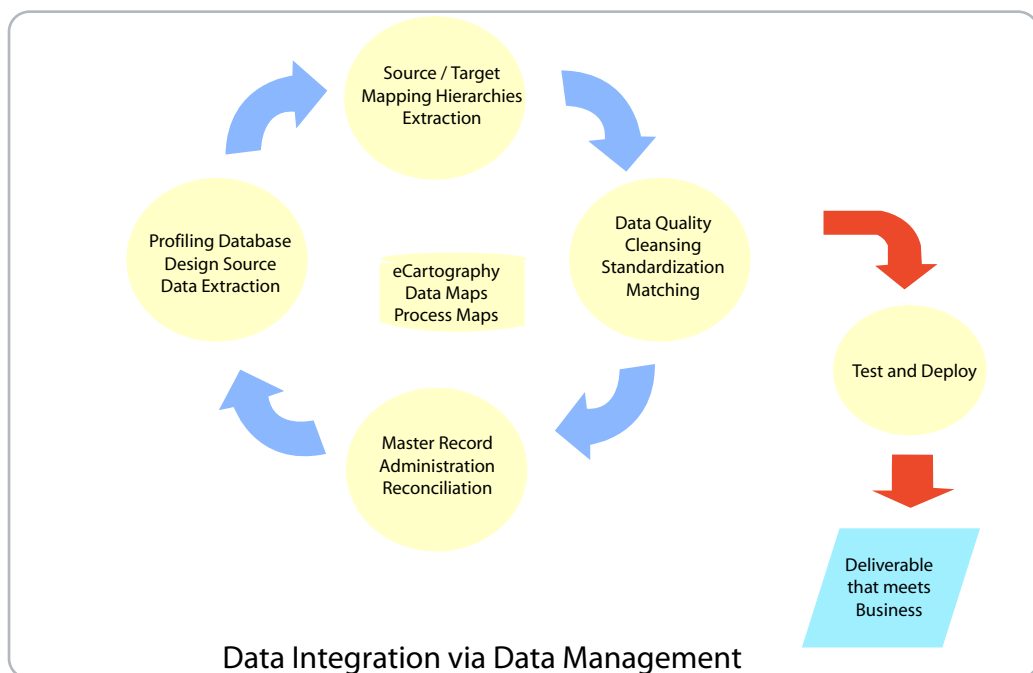
- » Surveying - Gathering technical metadata describing structure for a point in time.
- » Profiling - Used to measure tolerances (min., max., modes, etc.) for a point in time.
- » Pattern Analysis - Used to correct format issues.

- » Domain Analysis - Used to maintain and automatically update business domains and hierarchies for a point in time.
- » Relationship Analysis - Used to validate required business domain relationships for a point in time.

The client wanted an approach for implementing a predictive and automated process for data cleansing, quality and analysis. The premise is to use data profiling maps and associated process maps to predict the actions necessary to cleanse data or determine the further analysis steps on a column-by-column basis.

The client wanted the product to be written in Microsoft's Visual Studio 2005 C#, and also fully integrated into Microsoft's SSIS 2005 SQL Server SSIS, DB2 and using the .Net 2.0 Framework which can access most databases and file formats.

Client wanted to develop automation around the data processes like:



**Solution Classification :** Data Quality & Data Cleansing/ Data Scrubbing

## Distinguishing Features

We enhance all facets of the enterprise data including customer data, product data, business data and conformance of data to business rules. Based on our pioneering integrated technology framework our services deliver the following key benefits:

- » Standardization of data across the Enterprise
- » Elimination of duplicate records.
- » Customer addresses verification/correction.
- » Validation/Correction of data based on business rules.
- » Identification and elimination of nulls and invalid data.
- » Discovery of data patterns and frequency counts.
- » Identification of statistical distribution of data values.
- » Discovery of data integrity relationships and their conformity.

## Solution

The application now provides a superior data profiling application. This application supports the creation and maintenance of eCartography datamaps.

In addition the application predicts the actions (Domain, Relationship etc ...) required for individual columns and guide the user thru the analysis and correction. With the associated Microsoft SQL Server Integrated Services eCartography (SSIS) components, this product eliminates the need for external data profiling and data cleansing software provided such vendors as First Logic, Trillium, Evoke etc ... We have leveraged Microsoft's SQL Server Business Intelligence Development Studio and Microsoft Visual Studio 2005 C# to develop this application and associated SSIS Control flows and Data Flow Components.

We have also implemented the SSIS which is a series of data flow components that provides advanced data flow actions (pattern, matching, merging, cleansing etc ...) through an intuitive interface and targets at low to moderately skilled developers.

The components required are an eCartography datamap. The data flows comprise of the initial implementation of AMB Dataminers Self Healing Application Architecture. From a competitive perspective the eCartography data flows once integrated into Microsoft SQL Server Integrated Services eCartography (SSIS) provides the equivalent to the data quality/cleansing capabilities the leading DQ vendors. Also included is a custom control flow designed to enable scheduling and processing datamaps created in eCartography via the Microsoft SQL Server Integrated Services (SSIS).

Q3 set up a dedicated team consisting of a team lead and software developers with internal program managers to monitor and guide the client-vendor partnership. Q3 ensured a transparent and flexible relationship by putting the right combination of people and technology under detailed consultation with technical managers at client site. Semi-detailed specifications were provided by client's and Q3's system architects worked extensively with peers both at Q3 and at client site for building a state of the art data base structure and detailed functional and design specifications. Complete operational transparency was maintained by keeping communication of status updates and progress as a regular process. People at client would hear from the team lead, the program manager and the developers on a regular basis, with more detailed daily status reports indicating progress against milestones.

### Operating Environment

Operating Systems Supported	(Not Applicable)
Databases Supported	IBM DB2 Oracle 9i Oracle 10G Microsoft Access Sybase Microsoft SQL Server
Networks Supported	(Not Applicable)

Programming languages Supported

(Not Applicable)

Middleware Supported

(Not Applicable)

### Implementation, Training & Support

Training Description

Self-managed training - little or no assistance is required.

Support Description

Telephone support - calls can be placed directly to a support staff.

E-mail support - questions can be sent directly to a support staff for review.

### Web Version Information

Data security, availability and reliability details:

Security is a critical aspect of our entire infrastructure. The following are a few of the steps that we use to ensure security of our systems and data.

- Our main processing systems are completely isolated from the Internet physically.
- Firewalls and intrusion detection software offer another layer of protection against any kind of external hacker intrusions.
- Round the clock monitoring of our systems against any kind of hacker attacks.
- Access to systems is permitted only to authorized personnel.
- Support for secure communications such as secure FTP, encryption of data, SSL etc.
- All customer data is physically removed from our systems on completion of data integration services.

Service Level Agreement (SLA) offering:

No

Additional hardware or software requirements:

No

## Benefits

- » Our solution is low cost and does not require consulting services to operate.
- » Eliminates the need for specialty data quality vendors and expensive solutions
- » All Data maps and Process maps are stored in an open shared repository, available to all team members. Business Analyst, Project Managers, Developers, DBA's etc...
- » Leverages the .NET developer environment and application architecture.
- » Scalable option of repositories on multi-platform database.
- » Extension of client's development teams globally by taking advantage of Q3's highly flexible Software Development Life Cycle (SDLC) methodology.
- » Business model and culture at Q3 entails ownership of the product development process.
- » It is not only the team members who are dedicated to a client; it is the whole company management which works in complete tandem and efficiency to ensure that the relationship is seamless and successful.
- » The problem of customer data integration (CDI) is solved, the cost and time savings are enormous.
- » The advantage of PDM approach is to greatly reduce the time and effort to move and integrate data.
- » PDM is an effective data quality strategy that helps you better understand your business environment, allowing you to maximize profitability and reduce costly operational inefficiencies.