
Data Quality Management – Maximize Your CRM Investment Return

White Paper
November 2007



We make your data usable

Data Quality Management – Maximize Your CRM Investment Return

Table of Contents

Overview	2
Introduction	3
The importance of data quality to the CRM On Demand user	3
Automated data cleansing is today's best practice	3
How to ensure accuracy and usability of data using data cleansing technology	5
Step 1: Clean & standardize all data prior to the initial data load	5
Step 2: Stop bad data from entering the CRM system	7
Step 3: Check for ongoing data quality within the CRM system	8
Summary	10
About us	10
About the authors	11

Overview

This is the first in a series of three white papers focusing on the critical nature of data quality for Oracle/Siebel CRM On Demand customers. This paper addresses data quality management issues and practices specific to Oracle/Siebel CRM On Demand and discusses the leading edge technology now available to users. This technology is an industry best practice to ensure the ongoing availability of high quality data.

Data Quality Management – Maximize Your CRM Investment Return

Introduction

The most essential and fundamental key to any successful customer relationship management (CRM) investment is data quality. Even the most sophisticated system in the world will ultimately fail without *high quality data* residing within the system.

The importance of data quality to the CRM On Demand user

In today's business world where information is the ultimate currency, the higher quality the data, the more valuable and powerful the data is to those who possess it.

Good and accurate data is the most important factor in determining how successful a company is in its usage of the CRM application. A successful CRM strategy is primarily determined by the quality of data residing in the application. High quality data is essential to a company that manages its sales and marketing goals using a CRM system.

The Oracle/Siebel CRM On Demand user has strong analytical suites available at their fingertips, but these powerful tools are dependent on the accuracy and usability of the data. If poor quality data is either imported or resides in the system, then the robust analytical potential of Oracle/Siebel CRM On Demand is rendered less powerful, if not entirely neutralized. Businesses throw people and money at the troubles caused by poor quality data, but technology that solves the problem more effectively, and with fewer resources, is now available.

Automated data cleansing is today's best practice

Data cleansing technology automates the manual data cleansing and integration projects and frees up resources for other projects. An investment in data quality management software costs less than one employee (salary, benefits, office space, etc.), and its implementation actually supports organizational goals of increased productivity and minimized costs.

Older methods that addressed data quality issues, such as writing custom scripts or manually reading and cleansing data, are inefficient when compared to an automated data cleansing tool.

Data Quality Management – Maximize Your CRM Investment Return



Figure 1. The Evolution of Data Quality Methodology

Custom script writing is hampered by the need to test and debug the scripts. The more critical the data, the more relevant testing becomes, thus the process becomes very time intensive. Scripts can work on smaller, initial data sets, but as the data sets grown in volume and complexity, the scripts fail.

Meanwhile, organizations that spend time manually cleansing data often find that the costs to maintain data accuracy are always escalating and contain hidden expenses, such as the additional human errors that occur while performing repetitive, mundane tasks.

Until recently, these were the only complete methods available, so organizations have had to live with these inefficiencies.

Today's CRM resources, however, demand better methods, and automated data cleansing is keeping pace with this demand.

Accurate data is the main factor in ensuring that the users are making correct key company decisions using the records and reports within the CRM application. An impressive and detailed report needs to contain, first and foremost, correct and accurate data. Without high quality data it is also difficult for a company to achieve its goal of high user adoption. Users will eventually stop using the application if they encounter faulty, stale or inaccurate data.

For these reasons, automated data cleansing technology is the key to solving the data quality issues inherent to any CRM system.

Data Quality Management – Maximize Your CRM Investment Return

How to ensure accuracy and usability of data using data cleansing technology

There are three critical steps to a complete data quality solution for your Oracle/Siebel CRM On Demand system:

1. Clean and standardize all data prior to the initial data load;
2. Stop bad data from entering the CRM system;
3. Check for ongoing data quality within the CRM system.

Step 1: Clean and standardize all data prior to the initial data load to Oracle/Siebel CRM On Demand

Initial data load projects often get stalled during the data cleansing and standardization phase. The initial CRM data comes from a variety of sources and applications, and its successful migration is a complex data quality project that is often underestimated. Manually cleansing the sum total of all the data is neither effective nor efficient, and contributes unnecessary costs to your CRM investment. In particular, when manual or scripted cleansing is done by those without Oracle/Siebel certification, there's a heavy learning curve for the user, which results in additional wasted time and expense.

The Oracle/Siebel CRM On Demand user must verify quality, integrity and format of data to achieve a complete data quality solution during migration. If such caution is not exercised, there will be failed or partially imported records, depending on whether the field is required or not.

Some of the tasks necessary to perform Step 1 include:

- Verify that all required data fields exist (i.e., name, telephone, etc.)
- Verify that required and identifying fields contain data and are not blank
The identifying fields would either be the On Demand key fields or the external unique ID
- Verify that key fields have unique data
- Check referential integrity (i.e., verify that all child entities are pointing to existing parent entries)

Data Quality Management – Maximize Your CRM Investment Return

- Check owner assignment as specified by Oracle/Siebel CRM On Demand user sign-in ID
- Format date/time fields properly
- Format telephone field according to the format acceptable to Oracle/Siebel CRM On Demand
- Format all checkbox fields to have yes/no values (note: these can be Y, y, Yes, YES, yes)
- Format all picklist fields to have valid picklist values
- Map address fields to the appropriate address fields according to country
- Ensure valid country and state values; the state values need to be the two-letter state codes, both letters capitalized
- Protect USA zip codes with a leading digit of zero (i.e., Microsoft’s Excel will drop the “0” from a zip code like 02413)
- Find and merge duplicates prior to initial import into Oracle/Siebel CRM On Demand

Figure 2 illustrates how intelligent data corrections are automatically made prior to the initial data load.

	City	State	Country	ZIP
Before	East Barre	Vermont	US	5649
After	East Barre	VT	USA	05649
Before	Boston	Mass.	United States	2115
After	Boston	MA	USA	02115

Figure 2a. Cleansing and standardizing address data; protecting U.S. zip codes

	First Name	Last Name	Address 1	Address 2
Before	BECKY	SHELTON	Suite 110A	227 Grand Avenue
After	Becky	Shelton	227 Grand Ave	Ste 110A
Before	lou	balboa	339 W 1st St, Dept C	
After	Lou	Balboa	339 W 1st St	Dept C

Figure 2b. Cleansing and standardizing name data; mapping address fields

Data Quality Management – Maximize Your CRM Investment Return

For this all-important first step, Oracle has certified and partnered with services organizations that specialize in CRM data quality. An organization that has an existing working relationship with Oracle professional services will ensure a smooth and timely migration. In step with industry best practices, these organizations use technology to achieve the highest possible quality of data, thus increasing accuracy, reducing errors, and promoting the efficient use of an organization's resources.

Users of Oracle/Siebel CRM On Demand step up to the leading edge of data quality management by choosing a fully automated, easy-to-use software that has the capacity and flexibility to respond to a complex data migration project without taxing an organization's IT professionals for custom scripts or requiring a major time commitment of other staff for manual cleansing. Also, custom scripts typically are not engineered to scale either in sophistication or in volume, as your data quality needs grow.

Step 2: Stop bad data from entering the CRM system

Regular data imports into Oracle/Siebel CRM On Demand must be as accurate as the initial data load. Poor incoming data compromises target lists, market segmentation, and response metrics. Today's cleansing software identifies and cleans up duplicate, incomplete and erroneous information within the CRM system. It also scours incoming files to guard against duplication in the CRM *before* the new files are imported thus maintaining a very high level of data integrity.

The most advanced tools that will keep data clean and maximize the potential of your CRM should:

- Import a CSV file, standardize it, and de-dupe it against your CRM system, all in one step
- Verify, fix, enhance, and standardize messy data
- Append data from external commercial data sources
- Compare and correct data against a master directory

Data Quality Management – Maximize Your CRM Investment Return

Step 3: Check for ongoing data quality within the CRM system

There are opportunities for data corruption from the first minute it is collected, making ongoing data quality a never-ending mission. For example, some percentage of data often is entered manually and changed continually, making this third step critical in the long-term maintenance of high quality data for your Oracle/Siebel CRM On Demand system.

The best tools to ensure ongoing data quality can:

- Match tricky typographical errors, such as “Pharmaceutical” and “harmaceutic,” or “Mancy” and “Nancy” using sophisticated algorithms
- Correct for nicknames, such as “Becky” and “Rebecca”, or “Dick” and “Richard”
- Correct automatically city and state spellings
- Collect dozens of duplicate records and funnel the key information, field by field, into one best record, which the user approves before merging
- Verify and update zip codes with internally incorporated postal data
- Place Address 1 and Address 2 data fields accurately if they have been accidentally reversed or merged into one field
- Ensure that mailings and other customer-facing communications are professionally formatted (i.e., *JANE smith* becomes *Jane Smith*)
- Standardize the spelling of company names. (i.e., *Oracle Co* becomes *Oracle Corporation*)
- Provide a CRM user with advanced functionality without the need to write scripts or learn code
- Allow for preview and customization of changes to data at the field level before merging into the CRM
- Match intelligently – examine the root of the data and find matches like “The Sports Authority” and “Sports Authority”, or “7th Heaven” and “Seventh Heaven”

Data Quality Management – Maximize Your CRM Investment Return

Figure 3 shows how today's software specialized to the needs of CRM data and the maintenance of high quality data pursues the comparison between these two data records. The fields that the software has recognized to contain conflicting data in the Account Name, First Name, Last Name, and Address 1 fields.

	Account Name	First Name	Last Name	Address 1	City	State	ZIP
1.1	Massachusetts Life Insurance	William	Cortés	334b 15th St	Austin	TX	78728
1.2	Nassacusetts Life-Insurance	William	Cortes	b334 Fifteenth Street	Austin	TX	78728
Final	Massachusetts Life Insurance	William	Cortés	334b 15th St	Austin	TX	78728

Figure 3. Duplicate identification example

In the Account Name fields software functions have identified similarities between the misspellings and punctuation differences and have grouped these as potentially duplicate records. The Final Record reflects the correct account name with the standardized spelling applied.

In the First Name fields, the software identified a duplicate record, after comparing the name William imported with the number "1" instead of the letter "l" and an existing, correctly spelled version.

The Last Name fields retains the accent on the "e" in Cortés in the final record – again the duplicate had been identified. In the Address 1 fields, a tricky duplicated address is identified for the merge into a single record.

The Final Record represents the best combination of data from each field selected for comparison resulting in the highest quality record.

In-depth analysis of data at the field level represents a best practice when implementing or modifying a data quality management program.

Data Quality Management – Maximize Your CRM Investment Return

Summary

In order to leverage information within the CRM application, high data quality is a must. A CRM application with up-to-date, accurate and good data allows individual users and companies to manage all their customer activities in an efficient and effective manner.

The robust analytical suites in Oracle/Siebel CRM On Demand are dependent on the accuracy and usability of the data the system is processing.

Today's data cleansing technology produces the highest quality data that maximizes the functionality of your CRM application. Data cleansing technology automates data quality and integration projects, eliminates the need for time-consuming script writing or manual cleansing, and lowers the impact of human error. Additionally, employment of the technology frees up time and resources for other projects.

The successful partnership between today's data cleansing technology providers and customer relationship management (CRM) systems results in a successful data quality management plan and the highest standard of data quality. With this powerful combination, you can be confident your CRM system is operating to its fullest potential.

About us

About ActivePrime – ActivePrime is the world leader in Contextual Data Shaping™ technology. Designed specifically for the CRM marketplace, its solutions use cutting edge, patent-pending algorithms to deliver next generation data cleansing solutions that recognize and utilize the context of CRM data.

ActivePrime's CleanMove™ Turnkey Service cleanses, standardizes and formats data specifically for initial data import or migration into Oracle/Siebel CRM On Demand.

ActivePrime CleanCRM™ ensures that CRM systems consistently maintain the highest level of data quality, and keeps data clean for all ongoing imports.

ActivePrime is the first data quality vendor to certify and partner with Oracle/Siebel CRM On Demand.

Data Quality Management – Maximize Your CRM Investment Return

About Oracle – Oracle is the world's largest enterprise software company. For nearly three decades, Oracle has provided the software and services that enable organizations to get the most up-to-date and accurate information from their business systems. Oracle's partner ecosystem, managed by Oracle PartnerNetwork, is fundamental to Oracle's success. Oracle PartnerNetwork is a global network of more than 19,500 companies who deliver innovative software solutions based on Oracle's leading products. Oracle is a registered trademark of Oracle Corporation and/or its affiliates.

Contact us

To ensure a successful and ongoing implementation for your Oracle/Siebel CRM On Demand system, contact:

ActivePrime, Inc.

617-247-9908

www.activeprime.com

sales@activeprime.com

About the authors

Rosaline Gulati is an ActivePrime co-founder and President. She has worked with hundreds of companies on issues of deploying data quality solutions with CRM systems. In 2006, Gulati was selected to chair the MIT CIO Symposium. Gulati has an MBA from MIT Sloan, an MS Electrical Engineering (Technology Policy) from MIT and a BSEE from University of Illinois, Urbana-Champaign.

Noni Azhar is an Oracle CRM On Demand Certified Senior Consultant at Oracle. He has ten years' experience as a Consultant, Software Developer, Team Lead, and Project Manager across various industries such as CRM, Medical/Health, Professional Services, Online Collaborative Systems and Data Management. Azhar has more than five years of experience implementing CRM systems and has worked on more than 400 CRM application deployment/implementation projects. Azhar has an MS Computer Engineering (Management Information Systems) from Santa Clara University.