WHITE PAPER

High data quality in the CRM system

The proverbial icing on the cake.

The goal of introducing a Customer Relationship Management system is to optimize and stabilize the relationships with existing and future customers in the long-term. However, an intelligent CRM system on its own is not enough for this. Correct, complete, up-to-date and therefore high-quality customer data in the CRM system is just as important as the system itself. The key to a satisfactory customer relationship is therefore not only the CRM system but also high data quality.

Indications of poor customer data quality include:

- High return rate in mailshots as a result of incorrect or incomplete addresses
- Customer complaints about multiple delivery of the same advertising mail
- Incorrect forms of address and address lines, when e.g. Mr Karen Miller and Mrs Walter Smith receive post
- Lack of confidence of the company personnel in the stored data

If the company does not succeed in achieving and maintaining a high data quality in the system, the potential of the CRM system cannot be fully utilized: long-term customer retention and increased efficiency when working with customer data therefore remain theoretical.

Various use scenarios of a CRM system are considered in the following. The areas of focus are the relationship of the data quality and the consequences of sub-optimum data. Furthermore, a practical approach shows how companies can obtain high quality data in a new CRM system or an existing one and how this status quo can be preserved.

> Correct, complete, up-to-date and therefore high-quality customer data in the CRM system is just as important as the system itself.

UNISERV

In touch with your customers

The introduction of a CRM system is a key factor for the long-term success of the company. The awareness of this has been increasingly developing in the managerial levels of many companies over the past few years.

Irrespective of the supplier and the components of the CRM system used, the focus is always on customer orientation and the underlying service concept. If the introduction of a CRM system is considered from an economic perspective,

> Well-managed business and customer relationships have a variety of very positive effects on your own company.

it quickly becomes clear that relationship management is always associated with the desire for long-term business relationships and the economic security connected with this. The CRM system should also contribute to stabilizing the business contact in the long run.

Well-managed business and customer relationships, i.e., a stable network of relationships, have a variety of very positive effects on your own company.

HERE ARE A FEW EXAMPLES:

- A satisfied customer is prepared to recommend the supplier and his products through simple word-of-mouth propaganda.
- In a good long-standing relationship between the customer and the supplier, the customer may make suggestions for improving products. Changing market demands are detected in this way, there are opportunities to optimize the product and the performance.
- A satisfied customer is more tolerant of price increases.
- Reliable data also has a direct effect on the motivation of the company personnel. Provided with the correct information, the colleagues can concentrate on the requirements of the customers. The customers feel well cared for, contact remains positive on both sides.

Every pot has its lid: Ideal connection of CRM & DQ

Regardless of whether an analytical or operative CRM is implemented, correct data determines success or failure.

An analytical CRM is used to carry out evaluations in the sphere of Business Intelligence with all the possible customer data. A high data quality is indispensable, in order to be able to carry out appropriate analyses and to make the right strategic decisions in the long-term. In an operative CRM, it is important that the contact data of the customer is correct, so that marketing campaigns and the service offer reach their target, i.e. the customer.

Once a high data quality has been provided in the CRM system, customers are solicited only once in marketing campaigns and therefore not irritated unnecessarily. The postage costs are also optimized. Sales opportunities and forecast analyses are reliable. Customer service is more effective, since all the relevant information is available in the correct form for direct customer contact.

Against this background, it becomes evident that data quality, i.e. correct and duplicate-free data, is an important prerequisite for the so-called "Single View of Customer" or "Single Point of Truth", because only optimum data really allows all the data relating to a customer to be compressed into one data record. A comprehensive view of a customer is thereby made possible.

Correct and duplicate-free customer data is the essential prerequisite for the "Single View of Customer".

IN CONCRETE TERMS, DATA QUALITY MEANS:

- Correct address data, also in the international environment, so that written correspondence reaches the recipient.
- Active addresses if places or streets have been renamed
- Registered changes of address, so that correspondence reaches the recipient at the new residence
- Registration of mergers and other changes at companies
- Deletion or deactivation of contacts, for which customer relationships cannot be maintained, e.g. deaths.
- Unique data records, so that the customer master is duplicate-free and a customer is therefore actually included only once in the database

Data quality in the CRM: a how-to guide

Regardless of whether a completely new CRM system is to be put in place, the data quality of an existing CRM system is to be optimized or several systems are to be combined to form a single CRM system, the requisite high data quality can be achieved in three sub-processes:

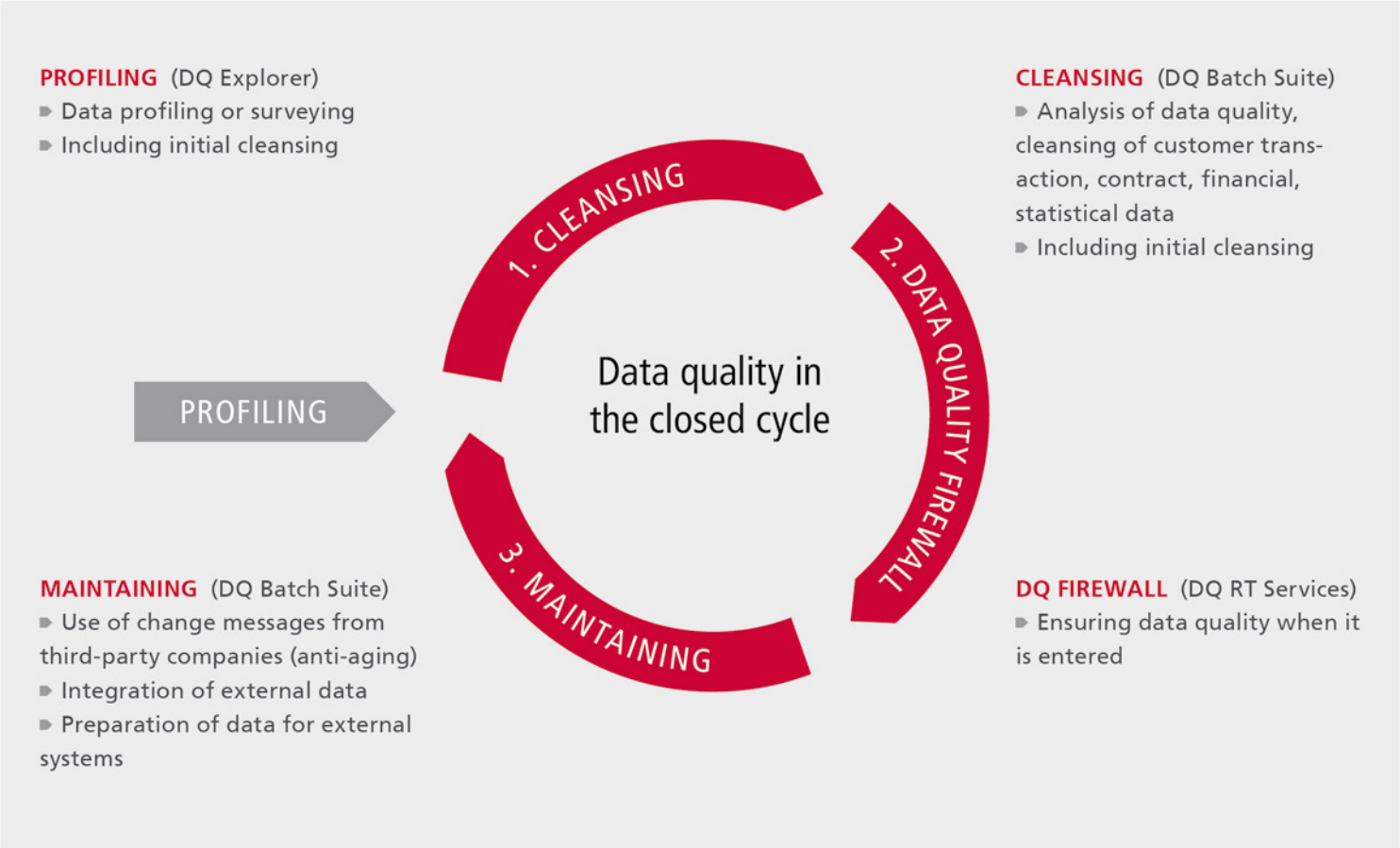
1. Initial data cleansing
2. "first time right" and mechanisms which intercept poor data quality when the data is entered or edited (Data Quality Firewall)
3. Use of data maintenance as a measure to preserve a high data quality standard

Data Quality Process

It is advisable to obtain an overview of the quality of the data in an upstream step, so that an initial cleansing oriented towards results is possible. This not only concerns correctly written addresses or duplicate data records but also learning about the structure of the data and checking the existing business rules. This step is typically implemented in a data quality audit. If a new CRM system is to be set up, the DQ Audit is a component of the DQ stream and therefore an important prerequisite for a successful migration. It acts in much the same way when additional source data systems are to be integrated in an existing CRM system.

Downstream monitoring is advisable for constant determination and verification of the status quo of the data quality. Compliance with the business rules can be automatically checked and critical threshold values specified here, in order to be able to carry out optimization measures in real-time. Such threshold values could also be key performance indicators (KPI), which provide information about the status quo of defined company goals.

> It makes no difference whether the CRM system is completely new or has been optimized or consolidated from different systems, the requisite high data quality can be achieved in three sub-processes!



The data quality process of Uniserv shows how the three sub-processes are connected with each other

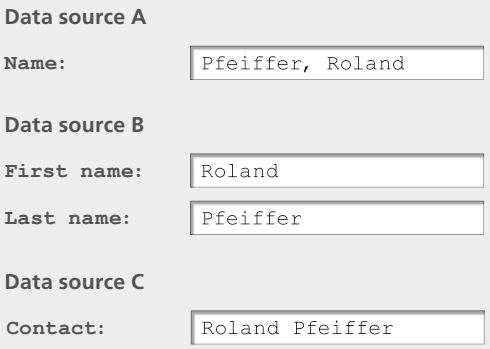
1. Initial data cleansing

The entire database is checked and cleansed in a batch run in the initial cleansing of the data. The number of different data sources or the countries which the data originates from are irrelevant here.

Typical sequence of the initial data cleansing

Standardized mapping: The field contents of different data sources are assigned to the same fields.

Example: The name of the contact person is in fields with different names in each data source. Ideally, the mapping should correspond to the mapping in the CRM system at this point. This ensures that the right data is loaded into the anticipated fields during the migration.



The data is converted to a standardized format.

Example: Standardized format for telephone numbers or dates.

The name components are analyzed.

Very complex name lines which either consist of several individuals or include the company name with the department and contact are analyzed. The analysis establishes whether the data is B2C data or B2B data. All the elements of the name line are also written to specially assigned fields, so that e.g. analyses of academic titles or legal forms of the company can be carried out.

A postal validation of the addresses is carried out.

The postcode, town, street and house number are checked for correctness irrespective of whether national or international addresses are concerned. If possible, missing address components are corrected and / or added. Officially renamed streets and places are automatically updated. PO box validation and bulk customer postcode validation are also available for certain countries.

Addresses of movers are updated.

Around eight million people change their place of residence in Germany each year. Only a very few of them actively advise of their new address. The data records of the individuals concerned can be updated initially or subsequently periodically by means of a relocation check over the entire database.

The addresses are converted to specific formats.

In certain countries, e.g., France, the address must be formatted according to the specifications of the national postal authorities, in order to be able to take advantage of postage rate optimization measures for the cheapest possible delivery options.

Stephane Petit	STEPHANE PETIT
Immeuble de corbusier Esc B	IMMEUBLE DE CORBUSIER
12 Route de Locmine	ESCALIER 8
56150 BAUD	12 ROUTE DE LOCMINE
	56150 BAUD

The addresses are enriched with additional information.

Depending on requirement, the addresses can be enriched with relevant information, such as geocoordinates or industry keys.

Rastatter Str. 13	+	Y-Koordinate +04889883
75179 Pforzheim		X-Koordinate +00866723

Duplicates are identified.

Duplicates, i.e. identical or similar data records, are identified by means of error-tolerant search algorithms. These search algorithms can be individually customized to the business rules. Uniserv offers standardized search algorithms for the B2C and B2B sector.

An assessment of the extent to which frontline employees have to manually decide whether the identified data records are duplicates or not can be made by means of a statement about the similarity of the discovered multiple data records.

Roland Pfeiffer		R. Pfeiffer
Rastatter Str. 13	=	Rastatterstrasse 31
75179 Pforzheim		75179 Forzheim

The "Golden Record" is formed.

The formation of a „Golden Record" is fundamental, particularly when data comes from a variety of sources which have further relevant contents attached in addition to the postal information. Data records of an identified duplicate group can be consolidated, i.e. information from the subsequent data records can be transferred to the header record of a group. Duplicate data records which are not eliminated can be marked with a deletion indicator.

Name	Street & No.	Zip code	City	Telephone	
Roland Pfeiffer	Rastatter Str. 13	75179	Pforzheim		Consolidation "Telephone"
R. Pfeiffer	Rastatterstr.		Forzheim	07231-936100	
		Result			
Roland Pfeiffer	Rastatter Str. 13	75179	Pforzheim	07231-936100	

When planning the initial filling of a CRM system, aspects relevant to migration should be considered in the initial cleansing.

This not only concerns e.g. the correct format of dates and telephone numbers but also formal things, such as the maximum length of street and place fields. If these relatively simple considerations are implemented in the initial cleansing, it is not necessary to implement complex transformation rules in the subsequent migration.

2. Data Quality Firewall: "first time right"

Specific standards for the initial creation of data records must be defined after transfer of the initially cleansed stored data. Only in this way is the obtained high data quality preserved.

A variety of real-time processing options present themselves here:

If a new data record is to be input in the system, it must first of all be checked whether it already exists. The creation of redundant information must be prevented. In much the same way as in the initial cleansing, an existing contact in the system is searched for very quickly by means of error-tolerant algorithms. The result of this search is either a list of similar data records, from which the appropriate entry is selected and edited, or the search is unsuccessful – in this case, a new record must be created.

Specific input rules must be complied with when a new data record is created in the CRM system. For example, street names should only be entered in the fields provided. A syntax check is possible in fields for telephone numbers. E-mail addresses can also be verified: here it is queried whether the domain and the user name actually exist. There is also the possibility of matching the stated postal address against reference data. This makes sense if e.g. the address was only given to a call centre by telephone and spelling mistakes or hearing errors falsify the data. If the information received over the telephone is incorrect or ambiguous, the employee can immediately ask for missing additional information such as the town district, in order to be able to transfer a postally correct address to the system. In the process, the address validation takes place either at the press of a button, during storage or fully automatically. Individual dialogs for address validation can also be offered as a matter of course.

Specific input rules must be complied with when a new data record is created in the CRM system.

The performance of the underlying technology is the key factor for the acceptance and use of these validation mechanisms. If these checks take too long, the functions will not be used – the Data Quality Firewall is by-passed. Since work is carried out under high time pressure, especially in call centres, the search for existing data records and the address validation must take place very quickly. The use of a rapid entry client enables automatic completion of the address components after input of the first letters or numbers.

Various alternatives are available for implementing these validation mechanisms:

There are basically two alternatives as far as postal validation is concerned. In the first place, the SaaS platform from Uniserv can be used. The advantage of this solution is that the installation and the regular update of the postal reference tables are no longer required. Invoicing is on a record-by-record basis, i.e. only what was actually validated is paid for.

On the other hand, the reference tables can also be kept locally and therefore addressed any number of times without incurring additional costs per request. The two solutions can also be mixed: licensed applications can be extended by services from the cloud in the hybrid model, e.g. the SaaS solution is used to validate countries, for which the purchase of reference tables is not worthwhile, whereas the addresses from the main countries are checked locally.

WebServices are the obvious choice for connecting a large number of CRM systems to the Uniserv address validation. These are generated at the touch of a button and administered by means of the Data Quality Real-Time Suite. The WebServices can then be directly integrated in the respective CRM system.
Uniserv makes specific DQ connectors available for certain CRM systems (SAP, Salesforce, Oracle Siebel, Microsoft Dynamics CRM and Update Seven). The advantage of this Plug-and-Play solution is that it comes with a graphical user interface for the address validation in the CRM system, which may have to be programmed in the case of the WebService solution.

A duplicate check in the cloud is only offered for Salesforce. Either WebServices or the relevant DQ Connector can be integrated in much the same way as described above. The search for potential duplicates has to be executed locally and not via the SaaS platform for all other CRM systems. The background for this is that the search for existing data records takes place in a code pool which is stored locally on the customer side. Amongst other things, this is for data security reasons. As in the case of the postal address validation, the duplicate check can be integrated via WebServices. Uniserv also offers DQ connectors for certain CRM systems.

> A hybrid model enables the SaaS solution to be used for specific countries whereas addresses from the main countries are checked locally.

The e-mail check and other checks (bank data, telephone numbers, relocations) can be integrated via WebServices, which, as already described, are easily generated and administered by means of the Data Quality Real-Time Suite. The e-mail check can only be accessed via the SaaS solution, the checks for correct bank data, telephone numbers and relocation processing operations can be integrated via the cloud as well as locally.

3. Data Maintenance: maintenance of the data quality standard

In spite of initial data cleansing and implemented DQ mechanisms, it is good policy to carry out a periodic check of the overall database.

There are around 40,000 official changes in the street and place directory in Germany each year. These concern the inclusion of new streets, local government reorganizations and renamings. Periodic matching against the current reference tables to take these changes into account and keep contact information up-to-date is therefore a good idea.
The several million changes of address in Germany each year should not be underestimated. Although this mainly applies to private households, not everyone who moves automatically notifies their banks, insurance companies or other business partners of their new address. Customer relationship management is a needlessly complicated challenge without up-to-date addresses.

Changes which have to be updated in the CRM system occur in the B2B sector as a result of mergers, insolvencies, renamings or a change in the legal form. The correct company name is a legal must when concluding contractual agreements.
Regular changes of name also occur in the B2C sector, e.g., as a result of marriages or divorces. If customers have died, the respective data records should be flagged at least.

> A periodic search for redundant data records in the overall database in a batch run is recommended.

Experience has shown that duplicate data records are created again and again in spite of the Data Quality Firewall. The firewall is by-passed, the matching data record is not found or changes have been made to the contact information which necessitate consolidation. A periodic search for redundant data records in a batch run in the overall database is therefore recommended.

The Data Quality Batch Suite can be used to implement data maintenance over the overall database. All the DQ checks can be configured via a standardized user interface by means of this powerful and highly parameterizable software suite. Several million data records per hour can be processed depending on the expansion level and check routine. The same business rules as in real-time processing and the Data Quality Firewall can be used in the search for duplicates.

CONCLUSION

Once the process steps described here have been implemented and a high data quality has been obtained, the potential of the CRM system can be fully utilized.

The evaluations in the analytical CRM are now on a sound basis. The data in the operative CRM permits a customer-oriented and efficient approach in all areas. Customer relationships are thereby strengthened in the long-term. The confidence of the company's employees in the quality of the data increases at the same time. One thing is clear: data quality in the CRM is the icing on the cake – make customer value management the secret to your success!

Data Quality Service Hub

Uniserv's Data Quality Services and Data Integration Services are assimilated in the modularly set-up Data Quality Service Hub.

The Uniserv Data Quality and Data Integration Services are designed to fit into the Data Quality Service Hub. This is a central data management platform for company data, in which the classic project scenarios of data management, such as data quality initiatives, data integration, data migration and consolidation, can be processed.

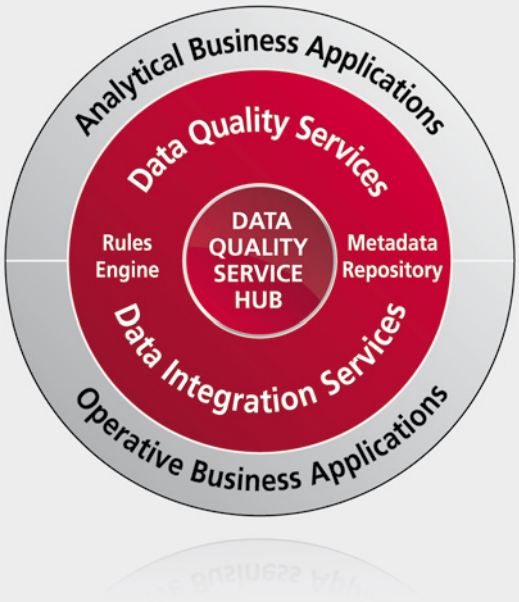
Together with Uniserv, you keep the quality of the company data at a high level over its entire period of use. The Data Quality Service Hub combines products which cover the complete Data Quality Cycle: from analysis of the database and an initial clean-up to the implementation of a Data Quality Firewall and regular cleansing of stored data as part of data maintenance. The Data Quality Service Hub consists of the following components:

DQ Explorer

This data profiling tool enables you to determine the status quo of your data quickly and interactively before transfer. You thereby obtain valuable information for the measures and priorities of your data quality management.

DQ Real-Time Suite

This acts as a firewall against contamination. The Data Quality Real-Time Suite secures the quality of data as it is input in interactive applications. It simplifies the integration of the DQ Real-Time Services (such as the postal validation or duplicate check) in applications



Data Integration Suite

This makes it easy to extract data from a great variety of sources, convert it into other formats and load it into destination systems – in batch processing and in real-time. An extensive set of connectors allows a wide range of data sources to be accessed and different data quality applications integrated via Plug&Play.

Data Quality Batch Suite

This is not a component of the Data Quality Service Hub but can be seamlessly integrated in a data stream. This is the Swiss Army knife of the data quality products for the validation, cleansing, consolidation and enrichment of data. The DQ Batch Suite undertakes the initial cleansing tasks, also for large and very large databases.

Data quality from the European market leader

Uniserv is the biggest specialized provider of data quality solutions in Europe with an internationally usable software portfolio and services for data management.

Data Management combines data quality assurance and data integration. With its solutions, Uniserv supports its customers in initiatives for data quality and projects for data integration, data migration and consolidation as well as data synchronisation, for example in the environment of CRM applications, eBusiness, direct and database marketing, CDI /MDM applications, data warehousing and business intelligence. With several thousand installations worldwide, Uniserv serves the expectation of a comprehensive solution for all business data throughout the entire data life cycle.

The company employs more than 110 people at its headquarters in Pforzheim, Germany and its branch in Paris, France. Its customers include numerous renowned companies from various industries throughout the world, such as Allianz, Amazon, Dell, Deutsche Bank, eBay, EDE- KA, E.ON, France Telecom, Johnson & Johnson, Lufthansa, OTTO, Siemens and Time Warner as well as TUI and Volkswagen.

For more information please visit www.uniserv.com.

UNISERV GmbH
Rastatter Str. 13, 75179 Pforzheim, Germany
T: +49 7231 936 - 0
F: +49 7231 936 - 3002
E: info@uniserv.com
www.uniserv.com

UNISERV