

A-Z

Terms from A to Z.

Glossary for Data Quality on Demand from Uniserv with specific technical terms for data quality and key terms which have a special meaning in the field of data quality.

A

Address format

Specification for the arrangement of addresses. The position of the individual address elements is different in each country.

Address validation

Data quality function by means of which the postal data of addresses is checked for correctness and completeness and corrected if required.

Algorithm

More or less complex set of rules which is used in Data Quality functions for the error-tolerant identification of individual data and optimization of complete files in addition to other special processes.

Audit

A special check of address files which Uniserv offers as part of its on-demand solution. This individual status analysis provides valuable information about the data quality.

B

Bank check

A data quality function by means of which the plausibility of the sort code and account number are checked.

Batch processing

The processing of tasks in a specified sequence. All the addresses stored in a file are processed in the Data Quality Center in batch mode. In contrast to single validation, processing cannot be executed immediately. Instead the complete file is gradually processed and can be downloaded from the Data Quality Center after processing has been completed.

Bereavement database

The data of up to 500,000 deceased persons is included in the bereavement database of Deutsche Post AG every year. The in-house data can be regularly cleansed by matching it against the database. Uniserv offers a service for this.

BIC

Abbreviation for Bank Identifier Code. This is an internationally standardized code by means of which each directly or indirectly participating partner can be uniquely identified worldwide. Is used by credit institutions in international transactions.

Business software

Collective term for software which is used in the company in the area of business management.

C

Check digit procedure

Procedure developed by banks/credit institutions for checking the validity of an account number. The check digit procedure reduces the danger of false entries and direct debits with incorrect account numbers.

Company matching

For company matching operations, you can select whether the matching should be carried out with or without matching of the contact person.

Complementary software solution

Complementary or supplementary software solutions are referred to at Uniserv if both Data Quality on Demand and the conventional software products (Data Quality on Premise) are used and complement each other.

Connectors

At Uniserv: special software which is used to establish connections between Uniserv applications and business software for data exchange. Connectors facilitate the link-up.

Country ID code

Abbreviation for the country.

Country-specific rules

Depending on the country, addresses are written according to country-specific rules. This country-specific formatting can be displayed in the single validation.

Credit card check

A data quality function from Uniserv which checks the plausibility of credit card numbers.

CRM

Abbreviation for Customer Relationship Management.

D**Data enrichment**

See Enrichment.

Data profiling

Data profiling refers to the largely automated process for the analysis of existing data stocks (e.g. in a database) using various analysis techniques.

Data Quality

Data quality makes a substantial contribution to optimizing corporate planning and controlling, opening up new business opportunities, minimizing risks and cutting costs. Good data quality is characterized by consistent, correct, up-to-date and duplicate-free customer and transaction data.

Data Quality account

Each user has his Data Quality account in the Data Quality Center. It summarizes his personal data, activities, status and quota.

Data Quality Report

A summary evaluation of the data quality during a file cleansing operation.

Data record

A structured quantity of information concerning a subject or an object, e.g. contact or address data.

Data Warehouse

A Data Warehouse is a central data collection whose content is composed of data from different sources. The data is provided by the data sources and loaded into the Data Warehouse. It is used for data analysis and as a tool for making business management decisions in the company.

Deduplication

Detection of multiple data records which refer to a person (or a company) and the merging of all data to create a single correct data record.

Demo client

The demo client offers a possible implementation for use of the Data Quality Center via Web Service. This can be used as a template to create your own client.

Direct mail refusers

Individuals who have entered their details in a Robinson list, thereby indicating that they do not wish to receive any advertising from companies with which they have no relationship.

Direct marketing

An advertising measure in the market communication of the marketing department which includes the direct address of the potential customer with an invitation for action, e.g. by letter or electronically, also via e-mail, is referred to as direct marketing.

DQ Connectors

See Connectors.

Duplicate group

If duplicates are found during a matching operation, they are internally aggregated to form duplicate groups. There is just one head duplicate in a duplicate group and one or more subsequent duplicates. The head duplicate is normally the address which is transferred to the 'cleansed' database, i.e. it is regarded as the 'only correct' address.

Duplicate type

The duplicate type is a flag for the found status of a duplicate. The following types can be found:

Single. The record has no correlation with another record. It does not belong to a duplicate group.

Head duplicate. The record is head of a duplicate group.

Subsequent duplicate. The record is a member of a duplicate group.

Ambiguous duplicate. The record is potentially a member of several different groups which have no correlation.

Duplicates

Several data records which refer to an object/company or a person and whose content completely or partly overlaps. Duplicates arise e.g. through hearing and recording errors when addresses are entered or when addresses from different business areas and departments are merged in a single database.

E**e-Commerce**

Commerce via Internet or webshop.

e-paper

Brochures, information material, etc. in digital form. A special reader or at least a PC with reader software is required for reading.

Enrichment

Appending additional information such as telephone numbers or geodata to data records.

ERP

ERP is the abbreviation for Enterprise Resource Planning. This refers to the deployment planning of the resources available in a company. Nowadays the software used to carry out the planning is normally meant by this.

F**File cleansing**

Data cleansing and data editing include various processes for removing and correcting data errors in databases or other information systems. The errors could arise from e.g. incorrect, out-of-date, redundant, inconsistent or incorrectly formatted data.

Form of address key

A key which ensures that the correct form of address is available for an address record. A key can have e.g. the following form: 1 = Ms 2 = Mr 3 = Sir or Madam

G**Geo-marketing**

Geo-marketing is a sub-discipline of marketing. It is devoted to geographical aspects in marketing and enhances traditional marketing with a spatial view.

Geocoordinates

The position of a point on the earth can be described by the geographic coordinates (latitude and longitude). If the coordinates of the addresses are known in the database, the address database can be turned into an efficient marketing instrument.

UNISERV offers knowledge bases (some free of charge) by means of which supplementary attributes can be assigned to the address records, including the geographic coordinates (longitude/latitude) for 13 countries in Europe. Additional reference data products (for Germany e.g. KGS (district/municipality key), freight, etc.) are available for many countries for the assignment of further attributes.

H**Head duplicate**

See "Duplicate type".

Household matching

There are various matching levels for each matching operation. These consist of the person and the household level for private addresses. A duplicate group is created for each individual person of a household for matching at person level, whereas all persons with the same/similar surnames and same/similar address are assigned to a common group for matching at household level. The persons living in a household can be summarized in this manner, whereby e.g. only one catalogue is sent to each household instead of to each member of the household.

I**IBAN**

IBAN stands for "International Bank Account Number". It is an international standardized notation for bank account numbers.

Integration

The data quality functions can not only be used in the Data Quality Center. Single validation and file cleansing can also be integrated in your in-house application as a Webservice. Ready-made Connectors for easy integration are already available for some applications, such as salesforce.com, Microsoft Dynamics or SAP.

K**KGS (District/municipality key)**

The territorial structure of the Federal Republic of Germany is hierarchical and is represented by the official KGS (district/municipality key). A statistical district is structured as the smallest unit according to: Land, Kreis, Regierungsbezirk, Gemeinde, Bezirk (Federal State, district, administrative district, municipality, area) and is described as a 16-figure district/municipality/street key. A municipality is divided into areas through assignment of the street identifier, incl. house number.

The statistical district can be assigned to the customer/prospect on the basis of the KGS.

L**M****Matching**

Data Quality function by means of which the contact data of customers, prospects, etc. is checked for redundancy in one or more databases. This not only prevents duplicates and enables contacts in B2B and B2C to be reliably identified but also allows different views of the address database to be mapped, e.g. in the consumer sector at the level of the person or household.

Microm

A company which specializes in target group and geo-marketing and provides the respective data.

Microtype

Provision of socio-regional data.

Mock-up

A (functional) model for testing or simulating certain components or processes.

N**Name analysis**

A data quality function in which the correct form of address is determined by means of elaborate algorithms and made available as a key.

Name line

The name line of a data record contains the first and second names as well as any additional name components, e.g. academic title. However, these address components can also be stored in individual fields.

O**On Demand**

The procurement of services and work - also computer services - as and when required.

Optimization (of data, data quality)

Measures for improving the data quality or for matching the quality to increased demands.

P**Person matching**

See Household matching.

Plug & Play

Insert, connect and forget.

Q**Quality (place, street)**

The quality is a measure and states how high the probability is that a piece of information is correct.

Quality management

As part of data management - the organized, instrumental processing of data according to quality criteria.

Quota

The quota provides information in the Data Quality account about how many validations can still be carried out.

R**Reference data**

Data from databases which has been accepted as correct and which can be used as a basis for comparison and correction.

Relocation database

The POSTADDRESS MOVE of Deutsche Post AG is the relocation file with the highest number of new relocation addresses per year in Germany. It only contains qualified relocation addresses. The addresses mainly originate directly from the more than 4 million private individuals and approx. 125,000 companies which relocate every year.

Result fields

The result fields are either written to the result file (in the case of file cleansing) or displayed directly on the screen (in the case of single validation). These are not only the fields of the processed data record but, depending on the application, also additional fields which append data to the data record or provide further information about the result of the processing.

Result file

The results of a data quality optimization operation are written into the result file and are available for further processing.

Robinson list

Database of the so-called direct mail refusers. These are people who do not wish to receive advertising mail from companies which are unknown to them. In Germany, such a database especially for postal items is maintained by the DDV (German Dialog Marketing Association).

S**Search algorithm**

Set of rules for finding defined objects, content, data and information or errors.

Service

Software service or service provision - in our context, a data quality function via Internet.

Single

See Duplicate type.

Single validation / online single validation

The interactive application of various data quality functions on a data record.

Street line

The street line of a data record contains the street name and the house number and possibly additional information depending on the country. However, these address components can also be stored in individual fields.

Subsequent duplicates

See duplicate type.

T**Task**

A task must be created and executed to carry out file cleansing in the Data Quality Center.

Task description

When a file cleansing task is created, a task description is generated which includes all the main parameters of the task in XML format. Task descriptions can be reused.

Telephone search

Data quality function for connecting personal data with telephone numbers or vice versa. Telephone data can be appended in addition to the simple search.

Town line

The town line of a data record contains the postcode and the town name as well as any additional location information depending on the country. However, these address components can also be stored in individual fields.

Transfer field

An input field which is passed to an output file without any changes.

U**Undeliverable**

Postal items which cannot be delivered because of incorrect addressing and are regarded as waste coverage in direct marketing.

Uniserv Connectors

Special software to connect Uniserv applications to business software. See also Connectors.

User status

Each user of the Data Quality Center has a defined status which specifies which services are available and how they are invoiced. Two statuses are currently defined: "Test" and "Productive".

V**Validation, postal**

Validation of the postal correctness of an address and its compliance with the country-specific form specifications.

W**Waste coverage**

Marketing measures and promotional messages are directed at target persons. Any messages which reach people other than the target persons or do not reach the target are summarized as waste coverage.

Web Service

Online service which the web technology uses to exchange services and data. A closely related term is SaaS. This is an abbreviation for "Software as a Service" and means the pure exchange of services which are normally generated by means of software.

XYZ