

# CROSS-DEPARTMENT SOLVING OF DATA QUALITY WEAKNESSES ORGANISATION – TOOL – PROCESS

Dr. Marcus Gebauer, Gerhard Knoke, Andrea Piro  
WestLB AG, Germany

**Abstract:** We show how a defined quality management process is modelled and efficiently used in a globally operating company for resolving data quality problems. This process is accompanied and supported by organisational measures and IT technology.

The organisational units rely on a central data quality management, which is responsible for the quality management process, in the solution of data quality problems.

**Key Words:** data quality, DQ problem, DQ organisation, central data quality management (CDQM), local data quality manager, measure, organisational change, process management, weakness

## 1. Introduction

WestLB is an internationally operating wholesale bank with branches, business units and subsidiaries all over the world. In terms of total assets, WestLB AG currently ranks number 30 in Europe and number 44 worldwide.

The Bank has set up a global IT and communication network in order to comply with the requirements resulting from its global presence. Managing the data and information flows of the heterogeneous systems and process landscapes involved is a challenging task. High data quality provides the basis for correct decisions and adequate risk management.

Hence, data quality management is of great importance. Data quality problems (DQ problems) often become apparent only at the end of a complete process chain, at the interface with clients, both internal and external clients. In a large Group, several organisational units are then affected by the same DQ problems. Trying to find out the causes of DQ problems over the entire process is very difficult. Hence, data quality management is also process management.

An individual solution often deals with the elimination of symptoms only, while the underlying causes continue to exist, not least as a result of activities involving several departments. A sustainable improvement of data quality is not achieved, and a multitude of different DQ problems appear in the systems of the Bank.

The analysis of the entire process chain and elimination of DQ problems necessitates the crossing of organisational borders and the cooperation of all organisational units involved in the process. The coordination of these tasks is achieved by central data quality management.

This paper deals with the process which permits the solution of DQ problems within the complex structures of the Bank in an intelligible und transparent form. This cannot be accomplished without the establishment of a corresponding DQ organisation. An IT tool provides an appropriate support for this process.

## 2. Approach

In the following sections, the organisational change and the new IT tool are described. By means of these features, an efficient reporting and solution process for DQ problems has been installed in the Bank.

### 2.1 Organisational Change

Figure 1 outlines the structure of the data quality organisation of the Bank. Central Data Quality Management (CDQM) supports the local Data Quality Managers (LDQM), who are responsible for the coordination of data quality measures to be implemented in their units. These include data cleansing measures, preventive measures for ensuring data quality, as well as data quality measurements. CDQM itself offers support to the LDQM and their business units in order to enable them to cope with these tasks. Thus, standardized, comparable DQ processes are guaranteed throughout the Bank.

CDQM is supported in the accomplishment of this task by several partner units which make their core expertise available. These include Group Organisation (Orga), Chief Information Office (CIO) and Mathematics/Operations Research (MOR).

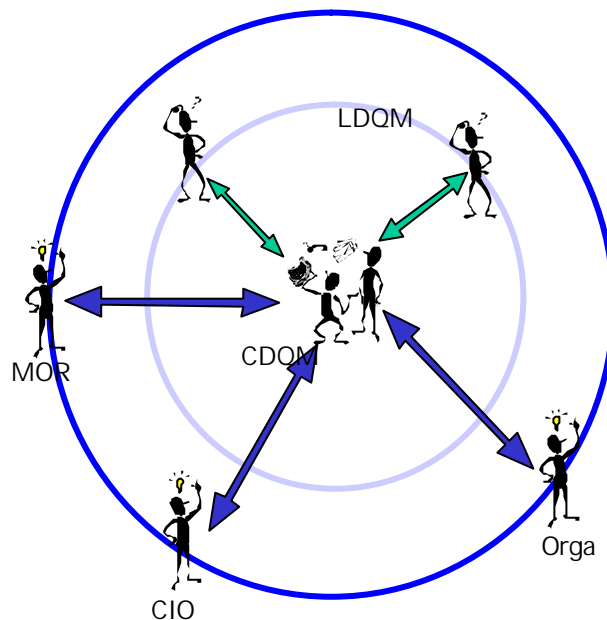


Figure 1: Outline of the organisational structure of holistic data quality management of the Bank

The LDQM should refer any questions concerning data quality to the CDQM which assumes a coordinating function in the solution of cross-department DQ problems. The process underlying the reporting and solution of DQ problems is illustrated in the following section "Tool".

## 2.2 Tool

It is of utmost importance that both the LDQM and all employees of the Bank can directly approach the CDQM and receive an immediate feedback for DQ problems reported. A hand-written "paper interface" as used in the past did not prove effective. It was not only impossible to obtain insight rapidly into the DQ problems of the Bank and to determine their causes, but it was also difficult to communicate the feedback to the reporting units. The following two sections deal with the requirements to be met by the tool and its integration into the solution process.

### 2.2.1 Requirements

The superordinate core requirements to be met by the tool result from the facts described in the introduction.

- Solution of cross-department DQ problems, support of worldwide structures of the Bank
- Utilization of synergies for problems occurring in several units of the Bank
- Coordination of the solution of problems by the CDQM
- Sustained improvement of data quality
- Evidence and transparency of the status of DQ problems
- Evidence and transparency of the major causes of poor data quality.

Figure 2 shows the communication channels required for solving a DQ problem and their link to a tool needed to support the reporting procedure for DQ problems efficiently. This tool must be able to show both the reported weaknesses and the measures to be taken. The tool implemented in the Bank is the WestLB quality data base.

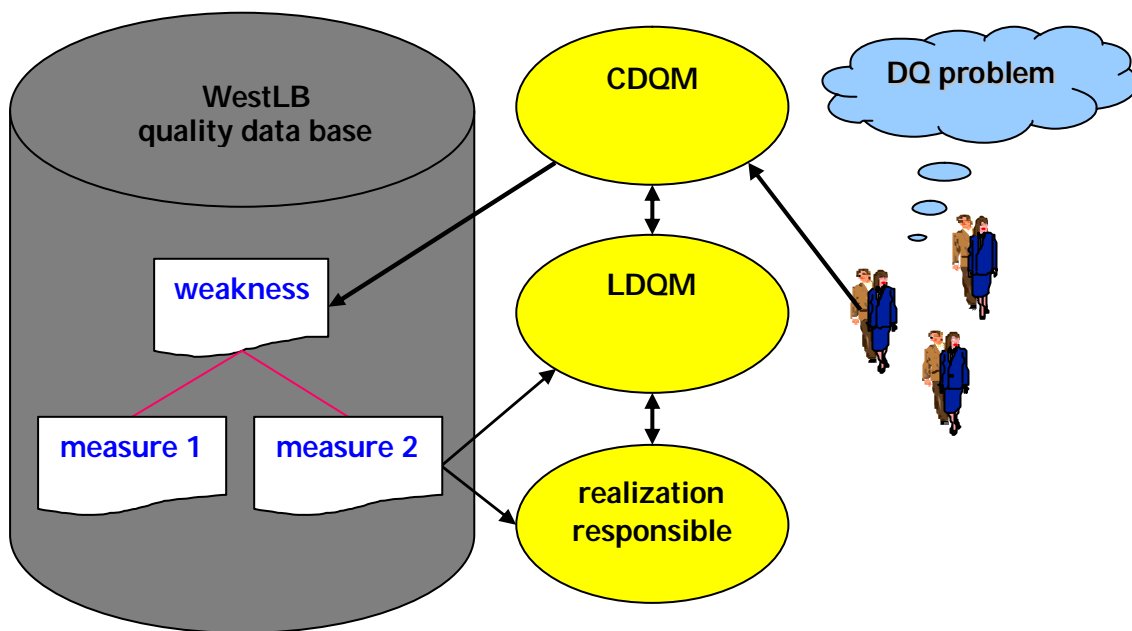


Figure 2: Illustration of the interaction among organisation, employees and the WestLB quality data base.

The process of resolving DQ problems requires the following functionalities of the WestLB quality data base:

- Provision of support for direct and rapid communication between CDQM, the LDQM and the persons responsible for realization.
- Automated feedback to the reporting unit to show the name of the employee of CDQM who registered the problem.
- Collection and storage of comments for documentation purposes.
- Written survey of the measures defined in order to create transparency of the solution process for all parties involved.
- Possibility of defining several measures for a weakness.
- Access to the updated data portfolio for all persons involved.
- Communication of the measures defined to the units responsible for realization.
- Automatic information of all persons involved on status changes of weaknesses and measures taken.
- Request for feedback to the responsible units concerning the status of the activities.
- Automatic reminder by standardized e-mail message if the deadline fixed for realization has expired or is exceeded.
- Flexible data base evaluation.
- Allocation of authorization profiles which regulate the access rights to the elements of the data base.
- Confidential treatment of data stored.

In order to meet the requirements, the application was implemented within the Lotus Notes e-mail system existing in the Bank. This ensures that the persons involved who work throughout the world can be reached at all times. In addition, it is possible to use the data base underlying the system for the defined requirements.

## 2.2.2 Functionality and Process

The functionality of the WestLB quality data base is shown by the integration into the overall process for resolving a DQ problem. This section provides a description of this process. For illustration purposes, important masks of the WestLB quality data base are shown.

The WestLB quality data base is available to CDQM, the LDQM and all reporting persons as a result of the implementation within the e-mail system. *Figure 3* shows an extract of the entry mask with a new message.

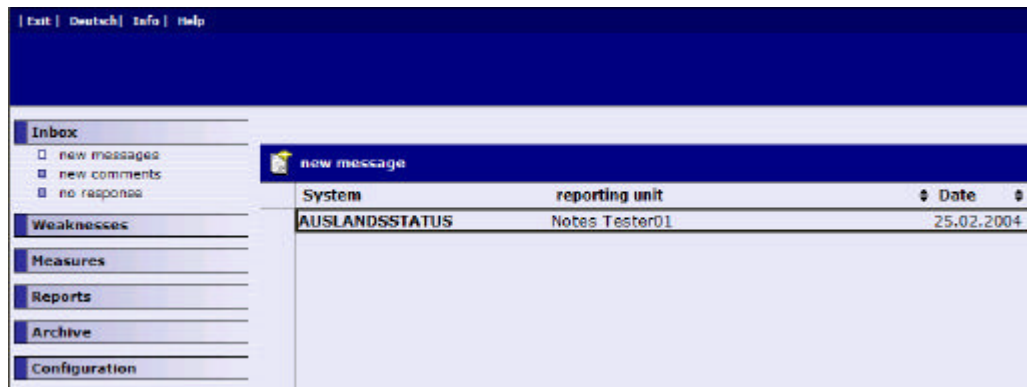


Figure 3: Entry mask of the WestLB quality data base

In the WestLB quality data base (Fig. 3), the views are shown as menu items on the left-hand side:

- Inbox
- Weaknesses
- Measures
- Archive
- Reports (CDQM only)
- Configuration (administrator only)

Under "Inbox" new messages concerning DQ problems and comments on existing measures are shown. However, comments in this view remain only until opened by the user. The sub-item "no response" is available to the CDQM only. It contains all measures in the check status which were not finally agreed with the LDQM.

The views "Inbox" through "Archives" are visible for all authorization profiles, whereas the item "Reports", in which total evaluations of all DQ problems can be prepared, is active for employees of CDQM only. "Configuration" is relevant only for the administrators, as they update their monitoring tables via this item.

The DQ problem solving process is modelled within a four-phase workflow (*Figure 4*), which is supported and maintained by the WestLB quality data base. After identification of the DQ problem (*inbox – phase 1*) the DQ problem and its causes are analysed (*analysis – phase 2*), which leads to the definition of the necessary measures. Processing (*processing – phase 3*) is followed by closure (*closure and archive – phase 4*) of the weakness after completion of all necessary activities. Afterwards, measures are defined which lead to the protection of the data quality achieved. The results of each individual step are communicated by e-mail to all units and persons involved.

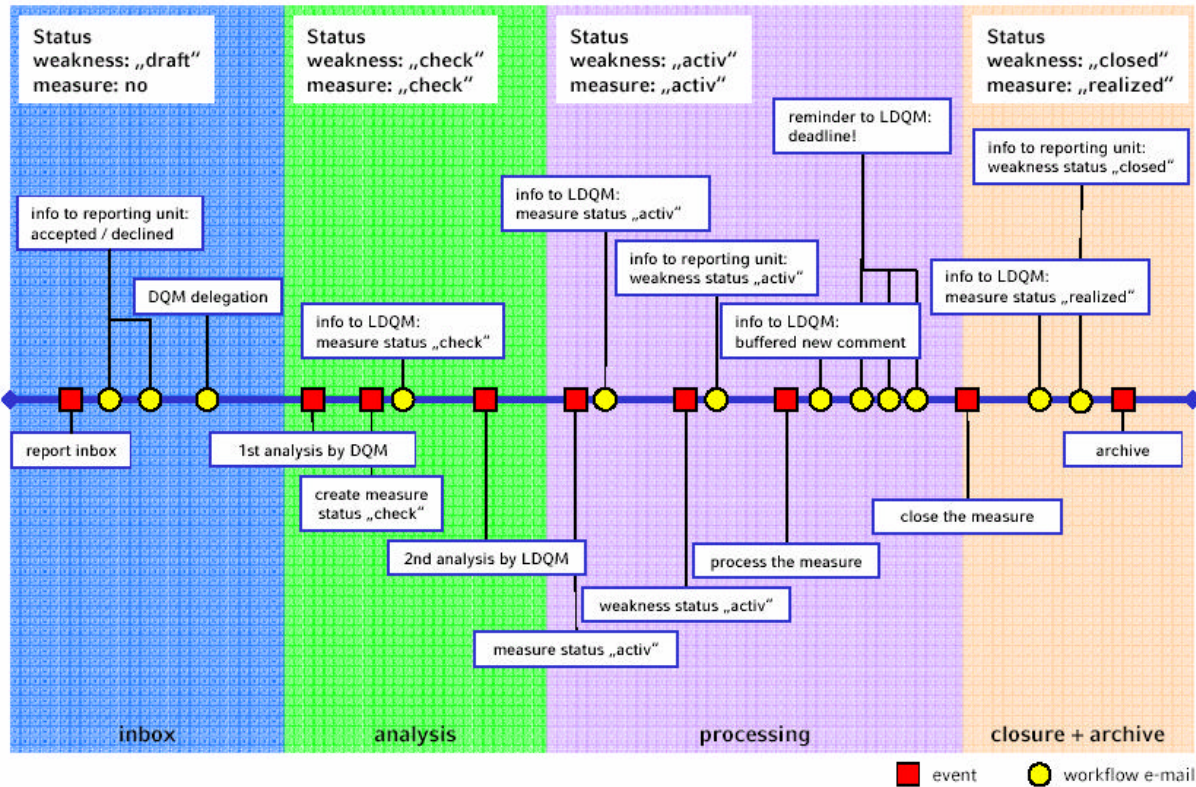


Figure 4: The four phases of solving DQ problems.

### Phase 1 – Inbox

An identified DQ problem is submitted via a message to the central data quality management. This DQ problem message is entered into the WestLB quality data base by means of a form (Figure 5) in which the problem is described. Simultaneously, initial references to systems and processes affected are made. The reporting person and his/her organisational unit are registered.

data quality weaknesses report	
<b>Identifikation</b>	
reporting person / BU:	Notes Tester01/D/WestLB_DEV
Telephone no.: / BU:	001
	2101
<b>further details</b>	
description:	problem description: eg. faulty calculation system

Figure 5: Reporting form in which the problem and relevant framework information is shown

The DQ problem report is shown in the item "Inbox" in the data base and can then be inspected. If there really is a DQ problem, the report is accepted by the CDQM and the reporting unit is automatically informed of the acceptance by e-mail. Within the CDQM, the DQ-problem (the weakness) is delegated to the responsible person.

After acceptance, the data base generates a "weakness" on the basis of the report (Figure 6), and an initial analysis is carried out by CDQM. At this point, the weakness is in the "draft" status, measures are not yet defined. In this initial analysis, the weakness is evaluated and its describing features determined (Figure 6):

- weaknesses number: consecutive number (year)/ current year at creation/ server name
- priority: valid values are *serious, material, minor*
- status (of the weakness): valid values are *draft, check, active, closed, declined*
- reporting person / BU: reporting person and his/her business unit
- data quality manager: responsible person of CDQM
- EDP system: systems related to the weakness/measures
- starting date: date of creation
- realized per: date of closure
- key word:
  - a. which process is firstly affected by the DQ problem
  - b. other describing key words
- description: short description of weakness and problems

The screenshot shows a web-based form titled "data quality weaknesses form". At the top, there are navigation buttons: "Close", "Edit", and "New Measure". The form is divided into two main sections: "identification" and "further details".

**identification**

number:	0000059 / 2004/D01PS	
priority:	serious	
status:	draft	
reporting person / BU:	Notes Tester01/D/WestLB_DEV	Ou No.:001
data quality manager:	Notes Tester01/D/WestLB_DEV	Ou No.:001
EDP systems:	CORONA	
starting date:	01.03.2004	16
realized per:		16

**further details**

keyword:	financing process - reporting
description:	problem description: eg. faulty calculation system
history / notice:	

At the bottom of the form, there is a link for "edit history".

Figure 6: An accepted weakness with describing features

Priority, which also serves for classifying the measures, is determined according to the effects already recognized and the frequency of their occurrence. As an efficient information platform, the WestLB quality data base offers the possibility to have recourse to similar or identical DQ problems and their possible solution.

## Phase 2 – Analysis

The solution of DQ problems is strongly determined by the processes and the environment in which the DQ problem occurs. Hence, in the analysis phase, the CDQM looks at the processes of the Bank. They are transparent for all employees via the Intranet in the WestLB process world. In addition to the IT systems with their data flows and responsibilities, the processes concerned are identified and entered into the WestLB quality data base as a key word (Figure 7). Figure 7 shows the classification of weaknesses according to the key words of the process world.

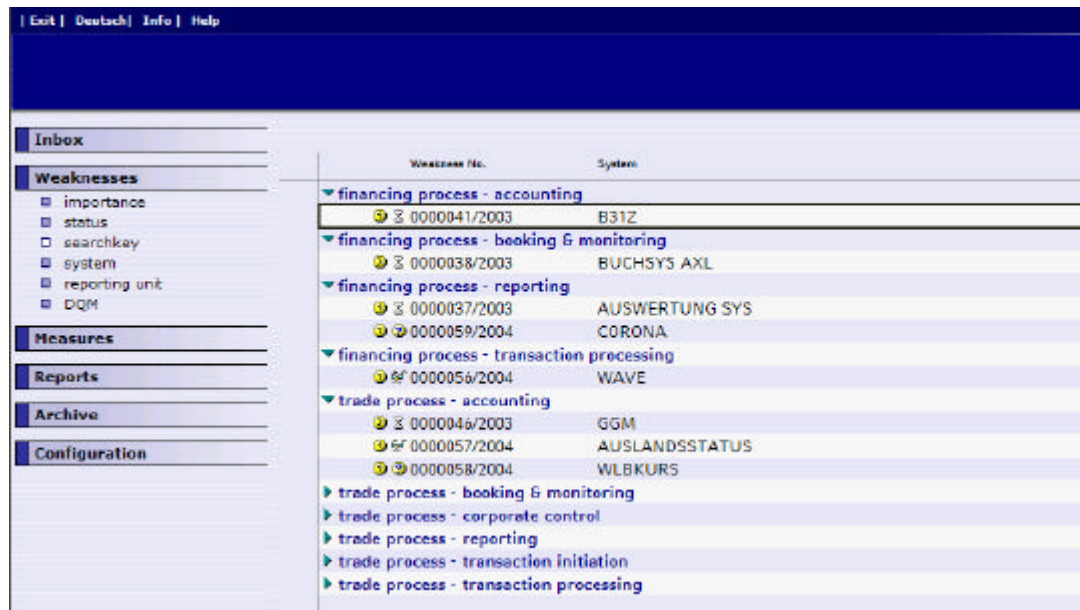


Figure 7: Illustration of weaknesses (right) according to various criteria (left). The searchkeys also show processes as concerned (e.g. "financing process")

The weakness is now in the "check" status. Initial measures are prepared as a draft and communicated to the responsible organisational units and their LDQM for verification. In cooperation with the responsible organisational units, the initial analysis is verified and detailed. The illustration of the context between causes and effects leads to the definition of the measures required to be taken. The measures are established in the "check" status. The LDQM are informed of the status.

## Phase 3 – Processing

Each measure is allocated to a weakness which can be accessed by a mouse click as shown in Figure 8 in the upper menu bar. Describing features are also specified for measures developed (Figure 8).

- number: number of the weakness
- status (of the measure): valid values are *check, active, realized*
- LDQM: LDQM of the responsible business unit
- responsible BU: business unit responsible for the measure
- problem category: valid values are *EDV (Engl. IT), human resources, organisation*
- problem: short description of the problem



- starting date: date of creation
- deadline: targeted deadline for the measure
- control date: new deadline, if measure is not realized at deadline
- finished at: date of realization
- description: short description of the measure

The screenshot shows a web-based 'Measure form' with a menu bar at the top containing 'Close', 'Save', 'Save & Close', 'Commit', 'Weakness', and 'Get Security'. The form is titled 'Measure form' and has a tabbed interface with the 'Identification' tab selected. The fields are as follows:

- number:** 000059/2004/D01PS
- Measure number:** 1
- status:** check
- DDQM:** Notes Tester02/ViewsLD\_DEV (BU No.: 001-10100, Tel No.: 2102)
- responsible BU:** Notes Tester03/ViewsLB\_DEV (BU No.: 001-10200, Tel No.: 1234)
- problem category:** EDV (radio buttons for Organisation and Personal are present)
- problem:** missing functions
- starting date:** 04.03.2004
- deadline:** 08.03.2004
- control date:** 08.03.2004
- finished at:** (empty)
- description:** implement a new calculation function

At the bottom, there is a 'history / notice:' section and a status bar with 'Weakness No.: 000059/2004', 'System: CORONA', and 'Keyword: financing process - reporting'.

Figure 8: A defined measure for weakness 59/2004/D01PS with related describing features.

After agreement of all measures to be taken, the measures and weakness are put to the "active" status and a date is fixed by which the measure should be realized ("deadline"). The measure is allocated to "Organisation", "Human Resources" or "Information Technology" (in *Figure 8* EDV). These are the three major category terms according to which CDQM classifies DQ problems. The responsible organisational units are supported by CDQM in the implementation of measures. During the implementation of measures the data base serves as the information and communication platform. Comments can be made which appear in the entry mask in the item "inbox – new comments" and can be called up by all persons involved. After the opening of a new comment, the advice is deleted in the "new comments" view. Within the measure, the comment is stored in a field (*Figure 8, bottom part*) and can no longer be edited after publication.

The reporting unit and LDQM are informed of the status of the weakness. The time manager of the data base issues a notice 14 days prior the scheduled deadline and reminds the responsible units automatically when the deadline is exceeded by 7 and 14 days.

#### Phase 4 – Closure and Archive

As soon as the measures are realized by the LDQM, CDQM is informed and the measure set to "realized" status. When all measures of a weakness are realized, the reporting unit is informed and requested to verify that the problem is solved. If the DQ problem is definitely solved, the weakness is closed by the CDQM and archived in the "closed" status (*Figure 9*). The relevant day of closure is entered in the measure and the weakness in the "realized per" field. The units concerned are informed of the completion of measures and weakness (e.g. *Figure 10*).

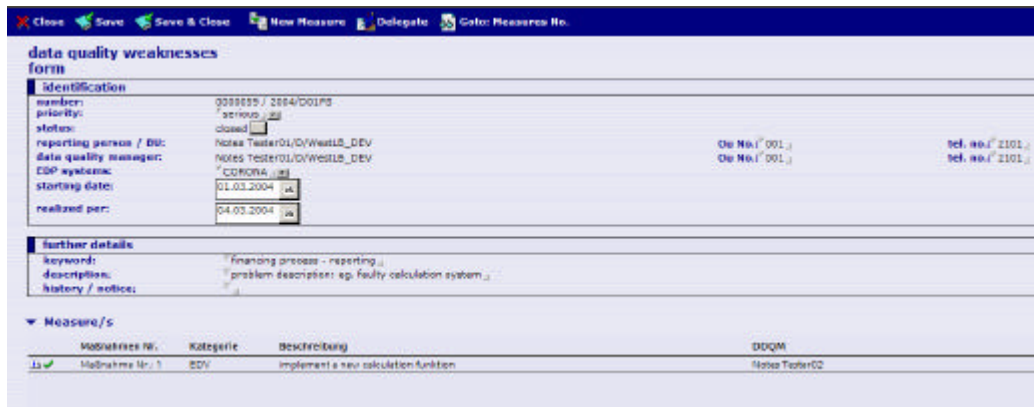
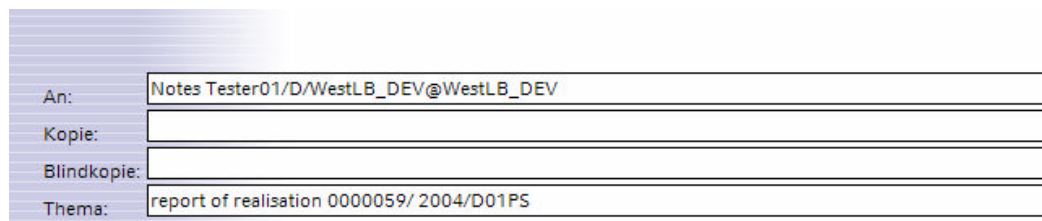


Figure 9: Entry of the weakness after realization of the measure. The realized measures are marked by a green hook.



Dear colleague,

we have been informed that the data quality problem which was reported is solved. Please check if you could agree too.

For further questions please feel free to contact us.

Best regards,  
your DQM Team


Please click on this document link. ---> 

Figure 10: The reporting unit is informed by e-mail of the completed weakness.

Closed weaknesses are recognizable in the "measures – by number" view by the green hook. In the "archive" view, the user can search for closed weaknesses by their closure date, weakness number and key words. A daily data base agent withdraws writing access from the user. Thereafter, closed weaknesses can only be reactivated by the Administrator.

The archived DQ problems enable the WestLB quality data base to serve as an information and know-how data base for both the Management and the LDQM, as the solutions provided always trigger ideas for the solution of new DQ problems.

### 3. Experiences and Results

In *Figure 11*, the number of DQ problems collected over the last nine years in the "Information Technology", "Organisation" and "Human Resources" categories is shown differentiated in an arbitrary scale. As from 1995, weaknesses have been analysed, categorized and systematically collected in an initial study. These and new weaknesses have since been solved.

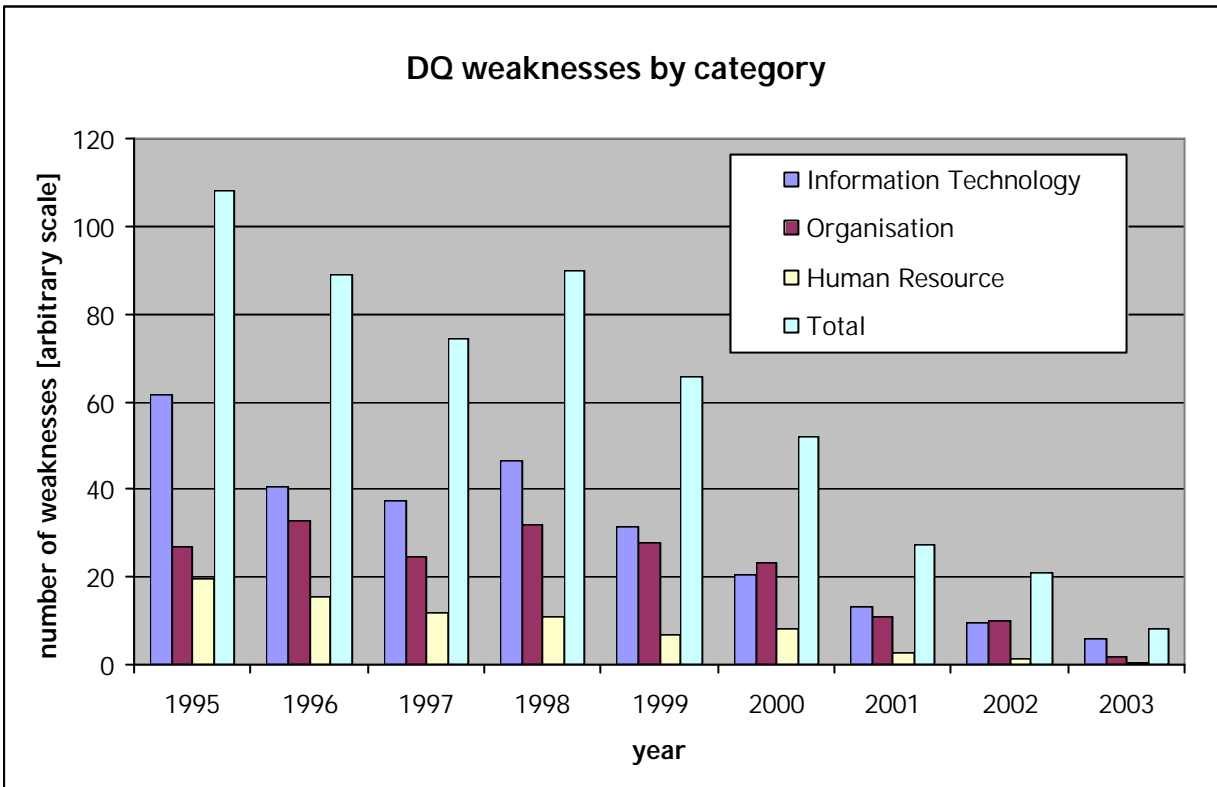


Figure 11: Development of the number of weaknesses over the last few years.

The implementation of a DQ organisation as described in Chapter 2.1 has proved to be a prerequisite for the elimination of weaknesses and implementation of the measures defined. The coordination of problem-solving with the LDQM has proved to be very effective.

However, this DQ organisation has always been questioned with regard to cost savings, so that the monetary benefits of data quality as accompanying measure must always be evidenced. The frequent exchange of LDQM practised by several units adversely affects the solution of problems and reduces the effectiveness of the process. If active data quality management is not considered a core task of the units, neither the tool nor the implemented process can develop their effectiveness.

Up to now a statistic about cost effectiveness of the DQ process doesn't exist. But an estimation of achieved synergies in projects dealing with the same topic can be shown. E.g. in one particular case (classification of customers in terms of their business sector) we were able to bring several projects and departments together to avoid doing the same things twice. So a (second) project with estimated costs of 1,000,000 € was no longer needed. We are convinced that living this process consistently will produce much more examples like this. Comparing this to the overall costs of 400,000 € per year for the whole process, employees payment and software and hardware it is very efficient.

#### **4. Conclusion and Outlook**

Evidence, transparency and support are mainstays in the solution of data quality problems. The organisational implementation of the LDQM and their integration into the reporting and solution process take account of this.

The process integrated high transparency in the solution of problems within the Bank as a result of the workflow in the WestLB quality data base. As the data base permits the evaluation of updated and archived DQ problems, the status of data quality is fully transparent for the Management and the persons concerned. The coordination of cross-department problems, which would be addressed simultaneously by several units without a uniform approach, offers large cost saving potential. Both the reporting units and the units responsible for solving the problem are supported and conflicts which might occur are mitigated by a neutral unit.

The WestLB quality data base provides the interaction and communication rules which support the organisational procedures of DQ improvement. The workflow system not only determines but also manages the process procedures of DQ improvement. Transparency is increased by the possibility to ascertain the current status of the process at any given time. Hence, it is possible to intervene flexibly in the process, to monitor its course and to react quickly to any new requirements. As a result of this new transparency, uncertainties can be eliminated at an early stage.

Excellent information transmission and communication in the process of DQ improvement is a major problem which can be solved by this concept. Experience gathered during the last few years has shown that the chosen path is the right way. However, it is also clear that the DQ organisation must be strengthened and confirmed at regular intervals. In addition to the implementation of the DQ Management Process described herein, the entire concept relies on regular information and sensitizing employees and Management alike to the need for sound data quality. The term quality must be deeply rooted in the corporate culture. If the term quality is established, a tool such as the WestLB quality data base can provide the necessary input for communication processes and contribute essentially to their improvement.

The WestLB quality data base will be enhanced in the next few years to include further reporting functionalities for increasing DQ transparency. The data base will be a mainstay for evaluating data quality in the company.