OPEN ACCESS
OPUS AUGSBURG

Universität Augsburg
Universitätsbibliothek

Metrics for measuring data quality - foundations for an economic data quality management

Bernd Heinrich, Marcus Kaiser, Mathias Klier

Angaben zur Veröffentlichung / Publication details:

Heinrich, Bernd, Marcus Kaiser, and Mathias Klier. 2007. "Metrics for measuring data quality - foundations for an economic data quality management." In Proceedings of the Second International Conference on Software and Data Technologies (ICSOFT 2007), July 22-25, 2007, Barcelona, Spain, volume 2, edited by Joaquim Filipe, Markus Helfert, and Boris Shishkov, 87–94. Setúbal: SciTePress. https://doi.org/10.5220/0001325600870094.

Nutzungsbedingungen / Terms of use: **CC BY-NC-ND 4.0**

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
CC-BY-NC-ND 4.0: Creative Commons: Namensnennung - Nicht kommerziell - Keine Bearbeitung
Weitere Informationen finden Sie unter: / For more information see:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

METRICS FOR MEASURING DATA QUALITY
Foundations for an Economic Data Quality Management

Bernd Heinrich, Marcus Kaiser and Mathias Klier
Department of Information Systems & Financial Engineering, University of Augsburg
Universitätsstr. 16, Augsburg, Germany

Keywords: Data Quality, Data Quality Management, Data Quality Metrics.

Abstract: The article develops metrics for an economic oriented management of data quality. Two data quality dimensions are focussed: consistency and timeliness. For deriving adequate metrics several requirements are stated (e. g. normalisation, cardinality, adaptivity, interpretability). Then the authors discuss existing approaches for measuring data quality and illustrate their weaknesses. Based upon these considerations, new metrics are developed for the data quality dimensions consistency and timeliness. These metrics are applied in practice and the results are illustrated in the case of a major German mobile services provider.

1 INTRODUCTION

In recent years data quality (DQ) has – due to an extended use of data warehouse systems, cooperative information systems (Cappiello et al., 2003) and a higher relevance of customer relationship management – gained more and more importance in science and practice. This refers to the fact that – for decision makers – the benefit of data depends heavily on completeness, correctness, consistency and timeliness. These properties are known as DQ dimensions (Wang et al., 1995). Many firms have problems to ensure DQ (Strong et al., 1997) and according to a study by Redman (Redman, 1998) "the total cost of poor data quality" is between 8 and 12 percent of their revenues. Moreover, an often cited survey by the DW Institute revealed poor DQ damaging US economy for more than 600 billions US-$ per year (The Data Warehousing Institute, 2002). Other statistics indicate that 41 percent of the data warehouse projects fail, mainly due to insufficient DQ (Meta Group, 1999). 67 percent of marketing managers think that the satisfaction of their customers suffers from poor DQ (SAS Institute, 2003). These figures illustrate impressively the relevance of DQ nowadays. The consequences of poor DQ are manifold: They range from worsening customer relationships and customer satisfaction by incorrect addressing of customers to a bad decision support of managers.

The growing relevance of DQ revealed the need for adequate measurement. Quantifying the current state of DQ (e. g. of a data base) is essential for planning DQ measures in an economic manner.

In the following we discuss how metrics for selected DQ dimensions can be developed with regard to two objectives:

a) Enabling the measurement of DQ
b) Analysing the consequences of DQ measures (e. g. data cleansing of customer's address data improving the quantified correctness)

The developed metrics were applied in cooperation with a major German mobile services provider. The objective of the project was to analyse the economic consequences of DQ measures in the case of campaign management. The following questions were relevant within the project:

- How can DQ be quantified and measured by means of metrics?
- How can DQ measures improve these metrics and what are the economic consequences?

The paper is organised as follows: The next section defines requirements on DQ metrics. In section 3 selected approaches are discussed. Section 4 develops new metrics for the DQ dimensions consistency and timeliness, and examines their advantages. A discussion of how the metric for timeliness was applied within the campaign management of a mobile services provider can be found in section 5. The

Heinrich B., Kaiser M. and Klier M. (2007).
METRICS FOR MEASURING DATA QUALITY - Foundations for an Economic Data Quality Management.
In *Proceedings of the Second International Conference on Software and Data Technologies - Volume ISDM/WsEHST/DC*, pages 87-94
DOI: 10.5220/0001325600870094
Copyright © SciTePress

last section sums up and reflects critically the results.

2 REQUIREMENTS ON DATA QUALITY METRICS

Applying an economic DQ management in practice, metrics are needed for quantifying DQ in order to answer questions like the following: Which measure improves DQ most and which one has the best benefit/costs ratio?

Figure 1 illustrates the closed loop of an economic oriented management of DQ. This loop can be influenced via DQ measures. Taking measures improves the current level of DQ (quantified by means of metrics). This leads to a corresponding economic benefit (e. g. enabling a more effective customer contact). Moreover, based on the level of DQ and taking into account benchmarks and thresholds, firms can decide on taking (further) measures or not. From an economic view, only those measures must be taken that are efficient with regard to costs and benefit (Campanella, 1999; Feigenbaum, 1991; Machowski & Dale, 1998; Shank & Govindarajan, 1994). E. g. given two measures having equal economic benefit, it is rational to choose the one with lower costs.

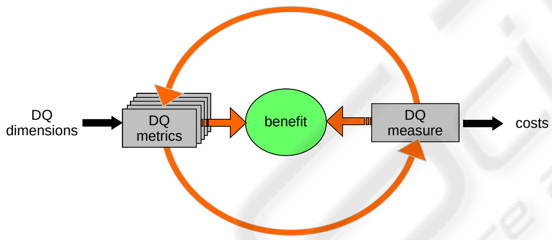

Figure 1: Data quality loop.

Therefore, this paper aims at quantifying the quality of a dataset by means of metrics for particular dimensions. The identification and classification of DQ dimensions is treated by many publications from both, a scientific and a practical point of view (English, 1999; Eppler, 2003; Helfert, 2002; Lee et al., 2002; Jarke & Vassiliou, 1997; Redman, 1996). In the following, we focus on two dimensions for illustrational purposes: consistency and timeliness. Within an economic oriented DQ management several requirements on DQ metrics can be stated for enabling a practical application (cp. Even & Shankaranarayanan, 2005; Hinrichs, 2002):

R1: [*Normalisation*] An adequate normalisation is necessary for assuring the results being interpretable and comparable.

R2: [*Cardinality*] For supporting economic evaluation of measures, we require cardinality (cp. White, 2006) of the metrics.

R3: [*Adaptivity*] For measuring DQ in a goal-oriented way, it is necessary, that the metrics can be adapted to a particular application.

R4: [*Ability of being aggregated*] In case of a relational database system, the metrics shall allow a flexible application. Therefore, it must be possible to measure DQ at the layer of attribute values, tupels, relations and the whole database. In addition, it must be possible to aggregate the quantified results on a given layer to the next higher layer.

R5: [*Interpretability*] Normalisation and cardinality are normally not sufficient in practical applications. In fact, the DQ metrics have to be interpretable, comprehensible and meaningful.

The next section reviews the literature considering the requirements listed above. Moreover it provides an overview over selected metrics.

3 LITERATURE ON MEASURING DATA QUALITY

The literature already provides several approaches for measuring DQ. They differ in the DQ dimensions taken into account and in the underlying measurement procedures (Wang et al., 1995). In the following, we briefly describe some selected approaches.

The AIM Quality (AIMQ) method for measuring DQ was developed at the Massachusetts Institute of Technology and consists of three elements (Lee et al., 2002): The first element is the product service performance model which arranges a given set of DQ dimensions in four quadrants. The DQ dimensions are distinguished by their measurability depending on whether the improvements can be assessed against a formal specification (e. g. completeness with regard to a database schema) or a subjective user's requirement (e. g. interpretability). On the other hand, a distinction is made between product quality (e. g. correctness) and service quality (e. g. timeliness). Based on this model, DQ is measured via the second element: A questionnaire for asking users about their estimation of DQ. The third element of the AIMQ method consists of two analysis techniques for interpreting the assessments. The first technique compares an organisation's DQ to a benchmark from a best-practices organisation. The second technique measures the distances between the assessments of different stakeholders.

Beyond novel contributions the AIMQ method can be criticised for measuring DQ based on the subjective estimation of DQ via a questionnaire. This approach prohibits an automated, objective and repeatable DQ measurement. Moreover, it provides no possibility to adapt this measurement to a particular scope (R3). But instead it combines subjective DQ estimations of several users who normally use data for different purposes.

The approach by (Helfert, 2002) distinguishes - based upon (Juran, 1999) - two quality factors: quality of design and quality of conformance (see also Heinrich & Helfert, 2003). Quality of design denotes the degree of correspondence between the users' requirements and the information system's specification (e. g. specified by means of data schemata). Helfert focuses on quality of conformance that represents the degree of correspondence between the specification and the information system. This determination is important within the context of measuring DQ: It separates the subjective estimation of the correspondence between the users' requirements and the specified data schemata from the measurement - which can be objectivised - of the correspondence between the specified data schemata and the existing data values. Helfert's main issue is the integration of DQ management into the meta data administration which shall enable an automated and tool-based DQ management. Thereby the DQ requirements have to be represented by means of a set of rules that is verified automatically for measuring DQ. However, Helfert does not propose any metrics. This is due to his goal of describing DQ management on a conceptual level.

Besides these scientific approaches two practical concepts by English and Redman shall be presented in the following. English describes the total quality data management method (English, 1999) that follows the concepts of total quality management. He introduces techniques for measuring quality of data schemata and architectures (of an information system), and quality of attribute values. Despite the fact that these techniques have been applied within several projects, a general, well-founded procedure for measuring DQ is missing. In contrast, Redman chooses a process oriented approach and combines measurement procedures for selected parts in an information flow with the concept of statistical quality control (Redman, 1996). He also does not present any particular metrics.

From a conceptual view, the approach by (Hinrichs, 2002) is very interesting, since he develops metrics for selected DQ dimensions in detail. His technique is promising, because it aims at an objective, goal-oriented measurement. Moreover this measurement is supposed to be automated. A closer look reveals that major problems come along when applying Hinrichs' metrics in practice, since they are hardly interpretable. This fact makes a justification of the metrics' results difficult (cp. requirement R5). E. g., some of the metrics proposed by Hinrichs - as the one for consistency - base on a quotient of the following form:

$$\frac{1}{(\mathit{result\ of\ the\ distance\ function})+1}$$

An example for such a distance function is $\sum_{s=1}^{|\Re|} r_s(w)$, where w denotes an attribute value within the information system. $\Re$ is a set of consistency rules (with $|\Re|$ as the number of set elements) that shall be applied to w. Each consistency rule $r_s \in \Re$ ($s = 1, 2, \ldots, |\Re|$) returns the value 0, if w fulfils the consistency rule, otherwise the rule returns the value 1:

$$r_s(w) := \begin{cases} 0 & \text{if } w \text{ fulfils the consistency rule } r_s \\ 1 & \text{else} \end{cases}$$

Thereby the distance function indicates how many consistency rules are violated by the attribute value w. In general, the distance function's value range is $[0; \infty]$. Thereby the value range of the metric (quotient) is limited to the interval $[0; 1]$. However, by building this quotient the values become hardly interpretable relating to (R1). Secondly, the value range $[0; 1]$ is normally not covered, because a value of 0 is resulting only if the value of the distance function is ∞ (e. g. the number of consistency rules violated by an attribute value has to be infinite). Moreover the metrics are hardly applicable within an economic-oriented DQ management, since both absolute and relative changes can not be interpreted. In addition, the required cardinality is not given (R2), a fact hindering economic planning and ex post evaluation of the efficiency of realised DQ measures.

Table 1 demonstrates this weakness: For improving the value of consistency from 0 to 0.5, the corresponding distance function has to be decreased from ∞ to 1. In contrast, an improvement from 0.5 to 1 needs only a reduction from 1 to 0. Summing up, it is not clear how an improvement of consistency (for example by 0.5) has to be interpreted.

Table 1: Improvement of the metric and necessary change of the distance function.

Improvement of the metric	Necessary change of the distance function
0.0 → 0.5	∞ → 1.0
0.5 → 1.0	1.0 → 0.0

Besides, we have a closer look at the DQ dimension timeliness. Timeliness refers to whether the values of attributes still correspond to the current state of their real world counterparts and whether they are out of date. Measuring timeliness does not necessarily require any real world test. For example, (Hinrichs, 2002) proposed the following quotient:

$$\frac{1}{(\text{mean attribute update time})\cdot(\text{age of attribute value})+1}$$

This quotient bears similar problems like the proposed metric for consistency. Indeed, the result tends to be right (related to the input factors taken into account). However, both interpretability (R5) – e. g. the result could be interpreted as a probability that the stored attribute value still corresponds to the current state in the real world – and cardinality (R2) are lacking.

In contrast, the metric defined for measuring timeliness by (Ballou et al., 1998) ($Timeliness = \{\max[(1-\frac{currency}{volatility}),0]\}^s$), can at least – by choosing the parameter $s = 1$ – be interpreted as a probability (when assuming equal distribution). But again, Ballou et al. focus on functional relations – a (probabilistic) interpretation of the resulting values is not provided.

Based on this literature reviews, we propose two approaches for the dimensions consistency and timeliness in the next section.

4 DEVELOPMENT OF DATA QUALITY METRICS

According to the requirement R4 (ability of being aggregated), the metrics presented in this section are defined on the layers of attribute values, tupels, relations and database. The requirement is fulfilled by constructing the metrics "bottom up", i. e. the metric on layer n+1 (e. g. timeliness on the layer of tupels) is based on the corresponding metric on layer n (e. g. timeliness on the layer of attribute values). Besides, all other requirements on metrics for measuring DQ defined above shall also be met.

First, we consider the dimension consistency: Consistency requires that a given dataset is free of internal contradictions. The validation bases on *logical* considerations, which are valid for the whole data and are represented by a set of rules $\mathfrak{R}$. That means, a dataset is consistent if it corresponds to $\mathfrak{R}$ vice versa. Some rules base on statistical correlations. In this case the validation bases on a certain significance level, i. e. the statistical correlations are not necessarily fulfilled completely for the whole dataset. Such rules are disregarded in the following.

The metric presented here provides the advantage of being interpretable. This is achieved by avoiding a quotient of the form showed above and ensuring cardinality. The results of the metrics (on the layers of relation and data base) indicate the percentage share of the dataset considered which is consistent with respect to the set of rules $\mathfrak{R}$. In contrast to other approaches, we do not prioritise certain rules or weight them on the layer of attribute values or tupels. Our approach only differentiates between either *consistent* or *not consistent*. This corresponds to the definition of consistency (given above) basing on *logical* considerations. Thereby the results are easier to interpret.

Initially we consider the layer of attribute values: Let w be an attribute value within the information system and $\mathfrak{R}$ a set of consistency rules with $|\mathfrak{R}|$ as the number of set elements that shall be applied to w. Each consistency rule $r_s \in \mathfrak{R}$ ($s = 1, 2, \ldots, |\mathfrak{R}|$) returns the value 0, if w fulfils the consistency rule, otherwise the rule returns the value 1:

$$r_s(w) := \begin{cases} 0 & \text{if } w \text{ fulfils the consistency rule } r_s \\ 1 & \text{else} \end{cases}$$

Using $r_s(w)$, the metric for consistency is defined as follows:

$$Q_{Cons.}(w,\mathfrak{R}) := \prod_{s=1}^{|\mathfrak{R}|}(1-r_s(w)) \qquad (1)$$

The resulting value of the metric is 1, if the attribute value fulfils all consistency rules defined in $\mathfrak{R}$ (i. e. $r_s(w) = 0 \ \forall \ r_s(w) \in \mathfrak{R}$). Otherwise the result is 0, if at least one of the rules specified in $\mathfrak{R}$ is violated. (i. e. $\exists r_s \in \mathfrak{R} : r_s(w) = 1$). Such consistency rules can be deducted from business rules or domain-specific functions, e. g. rules that check the value range of an attribute (e. g. $00600 \le$ *US zip code*, *US zip code* ≤ 99950, *US zip* *code* $\in$ {0, 1, …, 9}[5] or *marital status* $\in$ {"single", "married", "divorced", "widowed"}).

Now we consider the layer of tupels: Let T be a tupel and $\mathfrak{R}$ the set of consistency rules r_s ($s = 1, 2, \ldots, |\mathfrak{R}|$), that shall be applied to the tupel and the related attribute values. Analogue to the level of attribute values, the consistency of a tupel is defined as:

$$Q_{Cons.}(T,\mathfrak{R}) := \prod_{s=1}^{|\mathfrak{R}|}\left(1 - r_s(T)\right) \tag{2}$$

The results of formula (2) are influenced by rules related to single attribute values and rules related to several attribute values or the whole tupel. This ensures developing the metric "bottom-up", since the metric contains all rules related to attribute values. I. e. if the attribute values of a particular tupel are inconsistent with regard to the rules related to attribute values, this tupel can not be evaluated as consistent on the layer of tupels. Moreover, if the attributes of a particular tupel are consistent on the layer of attribute values, this tupel may remain consistent or become inconsistent on the layer of tupels. This decision is made according to the rules related to tupels.

In fact, a tupel is considered as consistent with respect to the set of rules $\mathfrak{R}$, if and only if all rules are fulfilled ($r_s(T) = 0 \quad \forall r_s \in \mathfrak{R}$). Otherwise $Q_{Cons.}(T, \mathfrak{R})$ is 0, regardless whether one or several rules are violated ($\exists r_s \in \mathfrak{R} : r_s(T) = 1$). Whereas consistency rules on the layer of attribute values are only related to a single attribute, consistency rules on the layer of tupels can be related to different attributes as e. g. (*current date* – *date of birth* < 14 years) $\Rightarrow$ (*marital status* = "single").

The next layer is the layer of relations: Let R be a non-empty relation and $\mathfrak{R}$ a set of rules referring to the attributes related. On the layer of relations the consistency of a relation R can be defined via the arithmetical mean of the consistency measurements for the tupels $T_j \in R$ ($j = 1, 2, \ldots, |T|$) as follows:

$$Q_{Cons.}(R,\mathfrak{R}) := \frac{\sum_{j=1}^{|T|} Q_{Cons.}(T_j,\mathfrak{R})}{|T|} \tag{3}$$

Finally, we consider the layer of data base: Assume D being a data base that can be represented as a disjoint decomposition of the relations R_k ($k = 1, 2, \ldots, |R|$). I. e., the whole database can be decomposed into pair wise non-overlapping relations R_k so that each attribute of the database goes along with one of the relations. Formally noted: $D = R_1 \cup R_2 \cup \ldots \cup R_{|R|}$ and $R_i \cap R_j = \varnothing \ \forall\ i \neq j$. Moreover, let $\mathfrak{R}$ be the set of rules for evaluating the consistency of the data base. In addition, $\mathfrak{R}_k$ ($k = 1, 2, \ldots, |R|$) is a disjoint decomposition of $\mathfrak{R}$ and all consistency rules $r_{k,s} \in \mathfrak{R}_k \subseteq \mathfrak{R}$ concern only attributes of the relation R_k. Then the consistency of the data base D with respect to the set of rules $\mathfrak{R}$ can be defined - based on the consistency of the relations R_k ($k = 1, 2, \ldots, |R|$) concerning the sets of rules $\mathfrak{R}_k$ - as follows:

$$Q_{Cons.}(D,\mathfrak{R}) := \frac{\sum_{k=1}^{|R|} Q_{Cons.}(R_k,\mathfrak{R}_k)\, g_k}{\sum_{k=1}^{|R|} g_k} \tag{4}$$

Whereas (Hinrichs, 2002) defines the consistency of a data base by means of an unweighted arithmetical mean, the weights $g_k \in [0; 1]$ allow to incorporate the relative importance of each relation depending on the given goal (R3). According to the approach of Hinrichs, relations that are not that much important for realising the goal are equally weighted to relations of high importance. In addition, the metric' results depends on the disjoint decomposition of the database into relations. This makes it difficult to evaluate objectively in the case of using unweighted arithmetical mean. E. g., a relation R_k with $k \neq 2$ is weighted relatively with $1/n$ when using the disjoint decomposition $\{R_1, R_2, R_3, \ldots, R_n\}$, whereas the same relation is only weighted with $1/(n+1)$ when using the disjoint decomposition $\{R_1, R_2', R_2'', R_3, \ldots, R_n\}$ with $R_2' \cup R_2'' = R_2$ and $R_2' \cap R_2'' = \varnothing$.

Now, consistency can be measured by using the metrics above in combination with corresponding SQL queries for verifying the consistency rules. The rules on the layers of attribute values and tupels can be generated by using value ranges, business rules and logical considerations.

After discussing consistency, we analyse timeliness in the following. As already mentioned above, timeliness refers to whether the values of attributes are out of date or not. The measurement uses probabilistic approaches for enabling an automated analysis. In this context, timeliness can be interpreted as the probability of an attribute value still corresponding to its real world counterpart (R5). In the following we assume the underlying attribute values' shelf-life being exponentially distributed.

The exponential distribution is a typical distribution for lifetime, which has proved its usefulness in quality management (especially for address data etc.). The density function $f(t)$ of an exponentially distributed random variable is described – depending on the decline rate $decline(A)$ of the attribute A – as follows:

$$f(t) = \begin{cases} decline(A) \cdot e^{-decline(A)\cdot t} & \text{if } t \geq 0 \\ 0 & \text{else} \end{cases} \quad (5)$$

Using this density function one can determine the probability of an attribute value losing its validity between t_1 and t_2. The surface limited by the density function within the interval $[t_1; t_2]$ represents this probability. The parameter $decline(A)$ is the decline rate indicating how many values of the attribute considered become out of date in average within one period of time. E. g. a value of $decline(A) = 0.2$ has to be interpreted as follows: averagely 20% of the attribute A's values lose their validity within one period of time. Based on that, the distribution function $F(T)$ of an exponentially distributed random variable indicates the probability of the attribute value considered being out-dated at T. I. e. the attribute value has become invalid before this moment. The distribution function is denoted as:

$$F(T) = \int_{-\infty}^{T} f(t)dt = \begin{cases} 1 - e^{-decline(A)\cdot T} & \text{if } T \geq 0 \\ 0 & \text{else} \end{cases} \quad (6)$$

Based on the distribution function $F(T)$, the probability of the attribute value being valid at T can be determined in the following way:

$$1 - F(T) = 1 - \left(1 - e^{-decline(A)\cdot T}\right) = e^{-decline(A)\cdot T} \quad (7)$$

Based on this equation, we define the metric on the layer of an attribute value. Thereby $age(w, A)$ denotes the age of the attribute value, which is calculated by means of two factors: the point of time when DQ is measured and the moment of data acquisition. The decline rate $decline(A)$ of attribute A's values can be determined statistically (see next section). The metric on the layer of an attribute value is therefore noted as:

$$Q_{Time.}(w, A) := e^{-decline(A)\cdot age(w,A)} \quad (8)$$

Thereby $Q_{Time}(w, A)$ denotes the probability that the attribute value is still valid. This interpretability is an advantage compared to existing approaches. Thereby the metric on the layer of an attribute value (5) fulfils the requirements normalisation (R1) and cardinality (R2). Figure 5 illustrates the results for different parameters $decline(A)$ depending on $age(w, A)$ graphically:

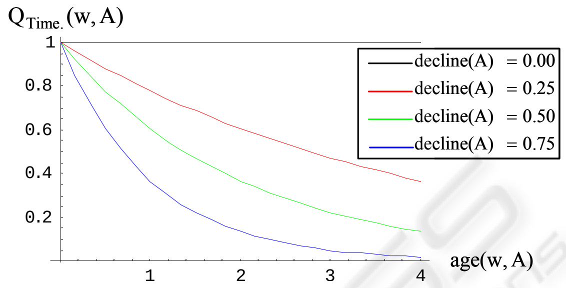

Figure 2: Value of the metric for selected values of decline over time.

For attributes as e. g. "date of birth" or "place of birth", that never change, we choose $decline(A) = 0$ resulting in a value of the metric equal to 1: $Q_{Time.}(w, A) = e^{-decline(A)\cdot age(w,A)} = e^{-0\cdot age(w,A)} = e^0 = 1$. Moreover, the metric is equal to 1 if an attribute value is acquired at the moment of measuring DQ – i. e. $age(w, A) = 0$: $Q_{Time.}(w, A) = e^{-decline(A)\cdot age(w,A)} = e^{-decline(A)\cdot 0} = e^0 = 1$ The re-collection of an attribute value is also considered as an update of an existing attribute value.

The metric on the layer of tupels is now developed based upon the metric on the layer of attribute values. Assume T to be a tupel with attribute values $T.A_1, T.A_2, \ldots, T.A_{|A|}$ for the attributes $A_1, A_2, \ldots, A_{|A|}$. Moreover, the relative importance of the attribute A_i with regard to timeliness is weighted with $g_i \in [0; 1]$. Consequently, the metric for timeliness on the layer of tupels – based upon (8) – can be written as:

$$Q_{Time.}(T, A_1, A_2, \ldots, A_{|A|}) := \frac{\sum_{i=1}^{|A|} Q_{Time.}(T.A_i, A_i) g_i}{\sum_{i=1}^{|A|} g_i} \quad (9)$$

On the layer of relations and database the metric can be defined referring to the metric on the layer of attribute values similarly to the metric for consistency (R4). The next section illustrates that the metrics are applicable in practice and meet the requirement adaptivity (R3).

5 APPLICATION OF THE METRIC FOR TIMELINESS

The practical application took place within the campaign management of a major German mobile services provider. Existing DQ problems prohibited a correct and individualised customer addressing in mailing campaigns. This fact led to lower campaign success rates.

Considering the campaign for a tariff option the metric for timeliness was applied as follows: Firstly, the relevant attributes and their relative importance (R3) within the campaign had to be determined. In the case of the mobile services provider's campaign the attributes "surname", "first name", "contact" and "current tariff" got the related weights of 0.9, 0.2, 0.8 and 1.0. Hence the customer's current tariff was of great relevance, since the offered tariff option could only be chosen by customers with particular tariffs. In comparison the correctness of the first name was less important. In the next step, the age of each attribute had to be specified automatically from the point of time when DQ was measured and the moment of data acquisition. Afterwards, the value of the metric for timeliness was calculated using decline rates for the particular attributes that were determined empirically or by means of statistics (see table 3 for an example).

Table 2: Determining timeliness by means of the developed metric (Example).

A_i	surname	first name	contact	current tariff
g_i	0.9	0.2	0.8	1.0
$age(T.A_i, A_i)$ [year]	0.5	0.5	1.5	0.5
$decline(A_i)$ [1/year]	0.02	0.00	0.20	0.40
$Q_{Time.}(T.A_i, A_i)$	0.99	1.00	0.74	0.82

The value of the metric on the layer of tupels is calculated via aggregation of the results on the level of attribute values, considering the weights g_i:

$$Q_{Time.}(T, A_1, \ldots, A_4) = \frac{0.99 \cdot 0.9 + 1 \cdot 0.2 + 0.74 \cdot 0.8 + 0.82 \cdot 1}{0.9 + 0.2 + 0.8 + 1} \approx 0.863 \cdot$$

Hence the resulting value of the metric for timeliness for the exemplary tupel is 86.3%, which means that the tupel is for the given application (promoting a tariff option) up to date at a level of 86.3%. The mobile services provider used these results in its campaign management. E. g. those customers (tupel) with a result below 20% were sorted out and did not receive any mailings. This is due to the fact that the results of former campaigns showed the segment of these customers being characterised by a success rate close to 0. Applying the metrics improved the efficiency of data quality measures and the measures could be evaluated economically. E. g., only those customer contact data with a value of the metric for timeliness below 50% were brought up to date by comparing these tupels to external data (e. g. to data bought from German Postal Service). This reduced the costs of acquiring addresses in a significant way. In addition, the effects on the metric's results and so the improvements of the success rates of the campaigns could be estimated. Thereby, the mobile services provider was able to predict the measures' benefits and compare them to the planned costs.

Besides these short examples for the metric's application resulting in both lower campaign's and measures costs, several DQ analyses were conducted for raising benefits.

By applying the metrics, the mobile services provider was able to establish a direct connection between the results of measuring DQ and the success rates of campaigns. Thereby the process for selecting customers was improved significantly for the campaigns of the mobile services provider, since campaign costs could be cut down, too. Moreover, the mobile services provider can take DQ measures more efficiently and it can estimate the economic benefit more accurately.

6 SUMMARY

The article analysed how DQ dimensions can be quantified in a goal-oriented and economic manner. The aim was to develop new metrics for the DQ dimensions consistency and timeliness. The metrics proposed allow an objective and automated measurement. In cooperation with a major German mobile services provider, the metrics were applied and they proved appropriate for practical problems. In contrast to existing approaches, the metrics were designed according to important requirements like interpretability or cardinality. They allow quantifying DQ and represent thereby the foundation for economic analyses. The effect of both input factors on DQ – as e. g. decline over time – and DQ measures can be analysed by comparing the realised DQ level (ex post) with the planned level (ex ante).

The authors are currently working on a model-based approach for the economic planning of DQ measures. For implementing such a model, adequate DQ metrics and measurement procedures are necessary. The approaches presented in this paper provide a basis for those purposes. Nevertheless, further metrics for other DQ dimensions should be developed.

Besides, the enhancement of the metric for timeliness in cases when the shelf-life can not be assumed as exponentially distributed is a topic for further research.

REFERENCES

Ballou, D. P., Wang, R. Y., Pazer, H., Tayi, G. K., 1998. Modeling information manufacturing systems to determine information product quality. In *Management Science*, 44 (4), 462–484.

Campanella, J., 1999. *Principles of quality cost*, ASQ Quality Press. Milwaukee, 3rd edition.

Cappiello, C., Francalanci, Ch., Pernici, B., Plebani, P., Scannapieco, M., 2003. Data Quality Assurance in Cooperative Information Systems: A multi-dimensional Quality Certificate. In Catarci, T. (edi.): *International Workshop on Data Quality in Cooperative Information Systems*. Siena, 64-70.

English, L., 1999. *Improving Data Warehouse and Business Information Quality*, Wiley. New York , 1st edition.

Eppler, M. J., 2003. *Managing Information Quality*, Springer. Berlin, 2nd edition.

Even, A., Shankaranarayanan, G., 2005. Value-Driven Data Quality Assessment. In *Proceedings of the 10th International Conference on Information Quality.* Cambridge.

Feigenbaum, A. V. 1991. *Total quality control*, McGraw-Hill Professional. New York, 4th edition.

The Data Warehousing Institute, 2002. *Data Quality and the Bottom Line: Achieving Business Success through a Commitment to High Quality Data*. Seattle.

Heinrich, B.; Helfert, H., 2003. Analyzing Data Quality Investments in CRM – a model based approach. In *Proceedings of the 8th International Conference on Information Quality.* Cambridge.

Helfert, M., 2002. *Proaktives Datenqualitätsmanagement in Data-Warehouse-Systemen - Qualitätsplanung und Qualitätslenkung*, Buchholtz, Volkhard, u. Thorsten Pöschel. Berlin 1st edition.

Hinrichs, H., 2002. *Datenqualitätsmanagement in Data Warehouse-Systemen,* Dissertation der Universität Oldenburg. Oldenburg 1st edition.

Jarke, M., Vassiliou, Y., 1997. Foundations of Data Warehouse Quality – A Review of the DWQ Project. In *Proceedings of the 2nd International Conference on Information Quality*. Cambridge.

Juran, J. M., 2000. How to think about Quality. In *Juran's Quality Handbook*, McGraw-Hill. New York, 5th edition.

Lee, Y. W., Strong, D. M., Kahn, B. K., Wang, R. Y., 2002. AIMQ: a methodology for information quality assessment. In *Information & Management*, 40, 133-146.

Machowski, F., Dale, B. G., 1998. Quality costing: An examination of knowledge, attitudes, and perceptions. In *Quality Management Journal*, 3 (5), 84-95.

Meta Group, 1999. Data Warehouse Scorecard. Meta Group, 1999.

Redman, T. C., 1996. *Data Quality for the Information Age*, Arctech House. Norwood, 1st edition.

Redman, T. C., 1998. The Impact of Poor Data Quality on the Typical Enterprise. In *Communications of the ACM*, 41 (2), 79-82.

SAS Institute, 2003. *European firms suffer from loss of profitability and low customer satisfaction caused by poor data quality*, Survey of the SAS Institute.

Shank, J. M.; Govindarajan, V., 1994. Measuring the cost of quality: A strategic cost management perspective. In *Journal of Cost Management*, 2 (8), 5-17.

Strong, D. M.Lee, Y. W., Wang R. Y., 1997. Data quality in context. In *Communications of the ACM*, 40 (5), 103-110.

Wang, R.Y., Storey, V.C., Firth, C.P., 1995. A Framework for analysis of data quality research. In *IEEE Transaction on Knowledge and Data Engineering*, 7 (4), 623-640

White, D. J., 2006. *Decision Theory*, Aldine Transaction.