# A PROCEDURE TO DEVELOP METRICS FOR CURRENCY AND ITS APPLICATION IN CRM

B. HEINRICH
Department of Information Systems, University of Innsbruck, Austria,
M. KAISER
Department of Information Systems Engineering & Financial Management, University of Augsburg, Germany
AND
M. KLIER
Department of Information Systems, University of Innsbruck, Austria.

---

Due to the importance of using up to date data in information systems, this paper analyzes how the data quality dimension currency can be quantified. Based on several requirements (e.g. normalization and interpretability) and a literature review, we design a procedure to develop probability based metrics for currency which can be adjusted to the specific characteristics of data attribute values. We evaluate the presented procedure with regard to the requirements and illustrate the applicability as well as its practical benefit. In cooperation with a major German mobile services provider, the procedure was applied in the field of campaign management in order to improve both success rates and profits.

Categories and Subject Descriptors: H.2.7 [**Database Management**] – Database Administration – Data warehouse and repository, J.1 [**Administrative Data Processing**] – Marketing, K.6.2 [**Management of Computing and Information Systems**] – System Management – Quality assurance

General Terms: Management, Measurement, Economics

Additional Key Words and Phrases: Data Quality, Information Quality, Metrics

Authors' addresses: B. Heinrich, Department of Information Systems, University of Innsbruck, Innsbruck, Austria, E-mail: bernd.heinrich@uibk.ac.at; M. Kaiser, Department of Information Systems Engineering & Financial Management, University of Augsburg, Augsburg, Germany, E-mail: marcus.kaiser@wiwi.uni-augsburg.de; M. Klier, Department of Information Systems, University of Innsbruck, Innsbruck, Austria, E-mail: mathias.klier@uibk.ac.at

---

## 1. INTRODUCTION

In recent years, the extended use of data warehouse systems and management support systems has given rise to a high relevance of data quality (DQ) issues in theory and practice [Cappiello et al. 2003]. This is due to the fact that – for decision makers – the benefit of data depends heavily on their completeness, correctness, and currency, respectively. Such properties are known as DQ dimensions [Wang et al. 1995]. Many studies deal with the costs and problems of poor DQ for companies and impressively illustrate the relevance of DQ, e.g. [Strong et al. 1997; Redman 1998; Meta Group 1999; SAS Institute 2003; Harris Interactive 2006]. A study by PWC brought up that only 34% of Chief Information Officers claim to be "very confident" in the quality of their data [PriceWaterhouseCoopers 2004]. According to a recent international survey on DQ, 75% of all respondents have made wrong decisions due to incorrect data. In addition, the respondents and their staff spend up to 30% of their working time checking the correctness of the provided data [Harris Interactive 2006].

Facing such problems, more and more firms want to improve DQ. However, it is essential to quantify the current state of DQ in order to plan DQ measures in an economic manner. Hence, procedures and metrics to quantify DQ are needed. This article focuses on the DQ dimension currency as several empirical investigations reveal that time aspects are very important in DQ management [Yu and Wang 2007; Klein and Callahan 2007]: We propose a procedure to develop metrics which shall enable quantification of the currency of attribute values stored in an information system.

Referring to the guidelines for conducting design science research defined by Hevner et al. [2004], we consider this procedure as our artifact and organize the paper as follows: After briefly discussing the relevance of the problem in this introduction, the next section describes the concept of Quality of Conformance and the DQ dimension currency. Moreover, six requirements are derived from literature (subsection 2.1). These requirements guide the process of searching for an adequate procedure to develop metrics for the DQ dimension currency: In subsection 2.2, an analysis reveals that existing approaches do not meet these requirements. The contribution of our research is therefore to close this gap by proposing a new approach. Hence, an innovative procedure based on probabilistic considerations is designed in section 3. An example (section 4) illustrates the individual steps of the procedure. Section 5 evaluates the procedure and its economic effects by means of a detailed real world example: The developed procedure was used within campaign management at a major German mobile services provider (MSP). The last section summarizes our findings from a managerial point of view and critically reflects on the results.

## 2. DATA QUALITY METRICS

Based on the current level of DQ and considering benchmarks and thresholds, firms have to decide whether to take quality measures (i.e., actions in terms of e.g. data cleansing, buying external address data etc.) or not. From an economic point of view, only measures which are efficient with regard to costs and benefit (cf. [Ballou and Pazer 1995; Feigenbaum 1991; Machowski and Dale 1998; Shank and Govindarajan 1994]) must be taken. The development of metrics is necessary to support this economic management of DQ and particularly to quantify the current level of DQ [Pipino et al. 2002; Campanella 1999; Heinrich et al. 2007]. In the following we describe the concept of quality that we focus on and motivate why we concentrate on the DQ dimension currency (which is also often named timeliness[1] in the literature).

In literature there are two different concepts and definitions of quality which also influence the quantification of quality: Quality of Design and Quality of Conformance [Helfert and Heinrich 2003; Juran 1998; Teboul 1991]. Quality of Design denotes the degree of correspondence between the users' requirements and the specification of the information system (which is for example specified by means of data schemata). In contrast, Quality of Conformance represents the degree of correspondence between the specification and the existing realization in information systems (for instance, data schemata vs. set of stored data values). Figure 1 illustrates these two concepts:

[1] The definition of *timely* is the following: done or occurring at a favourable or appropriate time. Since we intend to quantify, whether attribute values stored in an information system are still up to date, we use the term *currency* instead.

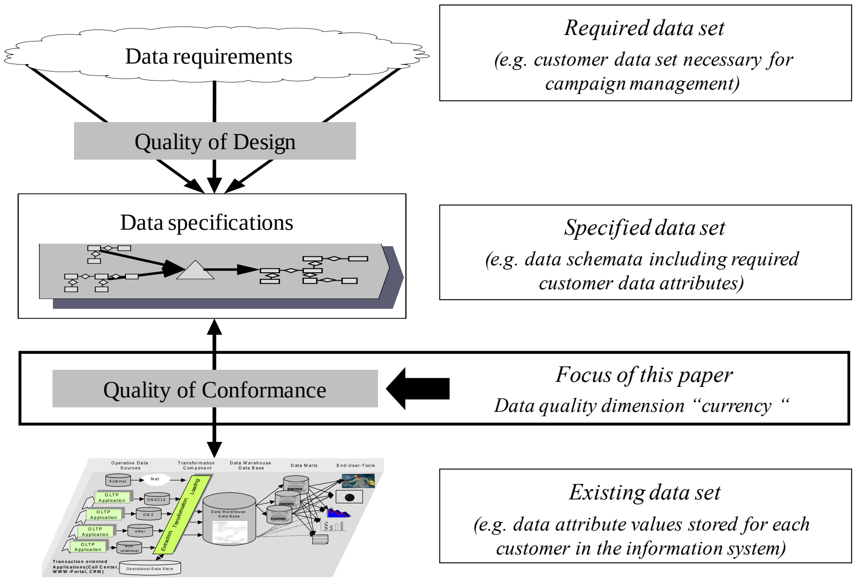


Fig. 1. Quality of Design vs. Quality of Conformance (cf. [Helfert and Heinrich 2003])

The distinction between Quality of Design and Quality of Conformance is also important in the context of quantifying DQ: It separates the (mostly subjective) analysis of the correspondence between the users' requirements and the specified data schemata (to check the quality of the information demand analysis) from the quantification of the correspondence between the specified data schemata and the existing data values – which is more objective. In the following we focus on Quality of Conformance, as metrics for quantifying this quality concept can be applied in many different situations and are therefore more reusable (because they are more independent from particular users' requirements in a specific business situation).

Considering the definition above, Quality of Conformance is mainly related to data values. According to Redman [1996], the DQ dimensions correctness, completeness, consistency and currency are most important in this context. These DQ dimensions have been discussed from both a scientific and a practical point of view in many publications, e.g. [Batini and Scannapieco 2006; English 1999; Eppler 2003; Jarke and Vassiliou 1997; Lee et al. 2002]. In many cases the main problem is not data being incomplete. Instead, it is more important to keep large sets of customer data, transaction data and contract data up to date. Hence, we focus on the DQ dimension currency and aim at quantifying the quality of data sets by designing a procedure to develop metrics for this dimension. In subsection 2.1, we derive six requirements for DQ metrics from literature. These requirements serve as evaluation criteria for existing metrics for currency in subsection 2.2. Moreover, they are guidelines for designing the procedure to develop a metric for currency in section 3.

## 2.1 Requirements for Data Quality Metrics

Many DQ metrics are designed on an ad hoc basis to solve specific, practical problems [Pipino et al. 2002]. Thus, they are often highly subjective [Cappiello et al. 2004]. To enable a scientific foundation and an evaluation of the metrics, we derive six normative requirements from literature. They are used to evaluate existing approaches for quantifying currency in subsection 2.2 and serve as guidelines for the search process

[Hevner et al. 2004, p. 88] when designing our own procedure in section 3. These requirements already proved to be useful when designing metrics for other DQ dimensions as e.g. correctness and completeness [Heinrich et al. 2007; Heinrich et al. 2008].

R 1. [*Normalization*] An adequate normalization is necessary to assure that the values of the metric are comparable (for instance, to compare different levels of DQ over time [Pipino et al. 2002]). Because of that, DQ metrics are often ratios with a value between 0 (perfectly bad) and 1 (perfectly good) [Pipino et al. 2002; Even and Shankaranarayanan 2007].

R 2. [*Interval scale*] To support both the monitoring of the DQ level over time and the economic evaluation of measures, we require the metrics to be interval scaled. This means, the difference between two levels of DQ must be meaningful. Thus, for instance, a difference of 0.2 between the values 0.7 and 0.9 and the values 0.4 and 0.6 of the metric means the same extent of improvement of DQ.

R 3. [*Interpretability*] Even and Shankaranarayanan demand the quantification to be "easy to interpret by business users" [Even and Shankaranarayanan 2007, p. 83]. For this reason, the values of the DQ metrics have to be comprehensible.

R 4. [*Aggregation*] In case of a relational data model, the metrics shall enable a flexible application. Therefore, it must be possible to quantify DQ on the level of attribute values, tuples, relations (especially views) and the whole database in a way that the values have consistent semantic interpretation on each level [Even and Shankaranarayanan 2007, p. 83]. In addition, the metrics must allow aggregation of the quantified values on a given level to the next higher level [Even and Shankaranarayanan 2007, p. 84]. For instance, the quantification of the correctness of a relation should be computed based on the values of the correctness of the tuples which are part of the relation. Moreover, the resulting values must have identical meaning as the DQ quantification on the level of tuples.

R 5. [*Adaptivity*] To quantify DQ in a goal-oriented way, the metrics need to be adaptable to the context of a particular application. If the metrics are not adapted, they should fold back to the non-adapted (impartial) quantification [Even and Shankaranarayanan 2007, p. 84].

R 6. [*Feasibility*] To ensure practicality, the metrics should be based on input parameters that are determinable. When defining metrics, methods to determine the input parameters shall be defined. If exact determination of input parameters is not possible or too cost-intensive, alternative rigorous methods (e.g. statistical) shall be proposed. From an economic point of view, it is also required that the quantification of DQ can be accomplished at a high level of automation.

These six requirements are used to evaluate existing approaches to quantify currency in the next subsection.

## 2.2 Literature Review

First, we define the term currency referring to corresponding literature. Afterwards, four approaches to quantify currency are analyzed whether they meet the six requirements.

In a first step, we analyze different definitions of the DQ dimension currency. Table I provides some selected definitions.

Table I. Selected definitions of this DQ dimension

| Reference | Term and Definition |
|---|---|
| [Ballou and Pazer 1985, p. 153] | Timeliness: “the recorded value is not out of date […]. A stored value, or any data item, that has become outdated is in error in that it differs from the current (correct) value.” |
| [Wang and Strong 1996, p. 32] | Timeliness: “The extent to which the age of the data is appropriate for the task at hand.” |
| [Redman 1996, p. 258] | Currency: “Currency refers to a degree to which a datum in question is up-to-date. A datum value is up-to-date if it is correct in spite of possible discrepancies caused by time-related changes to the correct value.” |
| [Hinrichs 2002] | Currency: “Property that the attributes or tuples respectively of a data product correspond to the current state of the discourse world, i.e. they are not out-dated” (own translation) |
| [Pipino et al. 2002, p. 212] | Timeliness: “the extent to which the data is sufficiently up-to-date for the task at hand” |
| [Batini and Scannapieco 2006, p. 29] | Timeliness: “Timeliness expresses how current data are for the task at hand.” |

In a Quality of Conformance context, the definition of Ballou and Pazer [1985] seems to be appropriate as it defines a DQ dimension which quantifies whether an attribute value stored in the information system is still up to date. This means that the attribute value, which was correct when it was stored, still corresponds to the current value of its real world counterpart at the instant when DQ is quantified. In other words, the attribute value has not become outdated (due to temporal decline).

In contrast to other dimensions like correctness, quantifying currency does not necessarily require a real world test. Instead, a metric for currency shall deliver an indication, not a verified statement under certainty, whether an attribute value has changed in the real world since its acquisition and storage within the system.

Based on this definition, we discuss the approaches by Hinrichs [2002], Ballou et al. [1998], Even and Shankaranarayanan [2007], and Heinrich et al. [2007] in detail and compare them to the six requirements, as these are – to the best of our knowledge – the only approaches which (1) design metrics for currency, (2) are based on a Quality of Conformance definition for the most part, and (3) are formally noted.

Hinrichs proposed the following quotient based on his definition of currency [Hinrichs 2002]:

$$Currency = \frac{1}{(mean\ attribute\ update\ \ frequency) \cdot (age\ of\ \ attribute\ value) + 1}$$

This formula shall quantify whether the current attribute value is outdated. The parameter *mean attribute update frequency* denotes how often the considered attribute value is updated on average within a certain period of time (e.g. 10 times per year). Regarding the input parameters, the quotient seems to return reasonable values: On the one hand, if the parameter *mean attribute update frequency* is 0 (i.e. the attribute value never becomes out of date), *currency* is 1 (attribute value is up to date). On the other hand, if the parameter *age of attribute value* is 0 (i.e. the attribute value is acquired at the instant of quantifying DQ), we get the same value. For higher values of *mean attribute update frequency* or *age of attribute value* the value of the metric approaches 0. This means that the (positive) indication (whether the attribute value is still corresponding to its real world counterpart) decreases. Hinrichs also provides formulas to aggregate the resulting values to higher levels, thereby his metric (partly) meets R 4. Moreover, the parameter *age of attribute value,* which is required to compute the value of the metric, can be extracted automatically (R 6) from the metadata in a database.

Despite these benefits, there are some shortcomings to consider which hinder economic planning and prohibit evaluating the efficiency of realized DQ measures a posteriori:

- The metric is normalized (cf. R 1), but the value range [0; 1] is generally not covered, because a value of 0 is only returned if *mean attribute update frequency* or *age of attribute value* respectively is ∞.
- The metric is hardly applicable within an economic DQ management, since the resulting values are not interval scaled (R 2). Therefore, neither absolute nor relative changes can be interpreted easily.

Table II illustrates the latter problem: To improve the value of *currency* from 0.0 to 0.5, the corresponding value of *mean attribute update frequency* multiplied with *age of attribute value* has to be decreased from ∞ to 1.0. In contrast, an improvement from 0.5 to 1.0 only requires a reduction from 1.0 to 0.0. Thus, a difference between two values of the metric (in our example 0.5) has no consistent meaning and the results of the metric are not interval scaled (cf. R 2).

Table II. Improvement of the metric and necessary change of parameters

| **Improvement of the metric** | **Necessary change of *mean attribute update frequency* multiplied by *age of attribute value*** |
|---|---|
| 0.0 → 0.5 | ∞ → 1.0 |
| 0.5 → 1.0 | 1.0 → 0.0 |

Furthermore, by building a quotient the results become hardly interpretable (cf. R 3) and cannot be interpreted by business users, for example (the value of the metric has no "unit"). Another limitation relates to the aggregation of the values of the metric from the level of relations to the level of the whole database: It is not possible to incorporate the relative importance of each relation depending on the given business context. Hence R 5 is not fully met.

Ballou et al. define the metric for timeliness as denoted below. In contrast, we refer to this dimension as currency (cf. above) and have moreover adapted the notation slightly (cf. [Ballou et al. 1998]):

$$Currency = \{\max[(1 - \frac{age\ of\ attribute\ value}{shelf\ life}), 0]\}^{s}$$

In contrast to Hinrichs [2002], the parameter *age of attribute value* is computed as follows: The time period between quantifying currency and acquiring the attribute value is added to the age of the attribute value at the instant of acquiring it. This corresponds to the age of the attribute value at the instant of quantifying DQ. The parameter *shelf life* is an indicator for the volatility of an attribute value. Thus, a relatively high *shelf life* results in a high currency and vice versa. The exponent *s* – which has to be assigned by experts – influences the extent to which a change of the quotient (*age of attribute value / shelf life*) affects the value of the metric. Thereby the computation can be adapted to the attribute considered and to the particular application to a certain extent (R 5). Moreover, the values of the metric are normalized to [0; 1] by means of the *max*-function (R 1).

However, it seems that the aim of Ballou et al. is to derive a mathematical function. Hence, they do not focus on getting values from the metric which are interpretable within an economic DQ management (cf. R 3). Similar to Hinrichs, in most cases the value of the metric has no "units". Indeed, the values are interpretable as the probability that the attribute value in the information system still corresponds to its real world counterpart only if $s = 1$. This case is equivalent to assuming a uniform distribution for the lifetime of an attribute value. However, a uniform distribution entails a fixed maximum lifetime and a constant (absolute) decline rate with regard to the initial value for a particular random variable. In the context of quantifying DQ, this means: For each considered attribute, a maximum lifetime which cannot be exceeded exists. This does not hold for many important attributes (e.g. *last name* or *date of birth*), as they possess neither a fixed maximum shelf life nor a constant (absolute) decline rate. For $s \neq 1$, the values of the metric cannot be regarded as probabilities relying upon common distribution functions and they are not interval scaled (R 2) any more. Table III illustrates the second shortcoming for $s = 2$. Again, an improvement of the metric by 0.5 has no consistent meaning. Therefore, it is obvious that such a metric cannot be adapted to all contexts (R 5).

Table III. Improvement of the metric and necessary change of parameters

| **Improvement of the metric** | **Necessary change of *age of attribute value* divided by *shelf life*** |
|---|---|
| 0.0 → 0.5 | 1.0 → 0.3 |
| 0.5 → 1.0 | 0.3 → 0.0 |

Ballou et al. do not propose formulas to aggregate the values of the metric to higher levels (R 4). Furthermore, the quantification of a particular attribute value's age can seldom be accomplished at a high level of automation (R 6). This is because the age of the attribute value at the instant of acquiring cannot be determined automatically from the metadata.

A utility-based approach for quantifying currency is presented by Even and Shankaranarayanan [2007]. The proposed metric is a function of the parameter *age of attribute value* taking values in [0; 1] and shall represent the user's utility resulting from the currency of the attribute value considered. Even and Shankaranarayanan [2007] present two examples in terms of utility functions which depend on the age of the attribute value.

The first utility function bases on the assumption that the user's utility resulting from the currency of an attribute value decreases exponentially with the age of the attribute value. Thereby, $\eta \in R^+$ represents the exponential decline factor (the larger $\eta$ is, the more rapid the utility declines with increasing *age of attribute value*):

$$Currency = \exp(-\eta \cdot age\ of\ \ attribute\, value)$$

When employing their second utility function, Even and Shankaranarayanan [2007] assume that an attribute value loses its utility completely when reaching the known maximum duration of validity $T_{max}$. Similar to the approach by Ballou et al. [1998], the exponent $s \in R^{+}$ can be used to influence the effect of the quotient (*age of attribute value* / $T_{\max}$) on the value of the metric:

$$Currency = \begin{cases} 1 - \left(age\ of\ attribute\ value / T_{\max}\right)^{s}, & \text{for} \quad 0 < age\ of\ attribute\ value < T_{\max}, T_{\max} \neq 0 \\ 0 & \text{for} \quad age\ of\ attribute\ value \geq T_{\max} \end{cases}$$

The values of the metrics are in both cases normalized to the interval [0; 1]. Besides, Even and Shankaranarayanan [2007] argue that the values can be interpreted as an (abstract) utility. However, an exact interpretation is not given (e.g., how shall a utility of 0.5 be interpreted?). In addition, the authors do not illustrate how adequate utility functions can be derived and interpreted, a weakness with respect to R 3. This also concerns R 2, as the results of the metric are interval scaled, if and only if the utility function quantifies a cardinal utility. From the authors' discussion on the two examples of utility functions it cannot be proven whether the utility functions are cardinal. This is because it is not assured that the values of the metric can be interpreted as an expression of preference strength. In contrast, R 4 is fulfilled, as Even and Shankaranarayanan [2007] provide formulas for aggregating the values of the metric to the levels of tuples, relations, database, and even several databases. Considering the two given examples of utility functions, they can be adapted to a particular business context to a certain extent by choosing the exponents $\eta$ and $s$ correspondingly. Moreover, Even and Shankaranarayanan [2007] argue that several utility functions can be used when quantifying DQ. In this respect, R 5 is met. However, a utility function must also be adapted to users' characteristics. Hence an automatic and objective quantification of currency is not possible (especially due to the fact that no details are given how such utility functions and the included parameters can be determined). Therefore, R 6 is only partly met.

The fourth approach presented in Heinrich et al. [2007] suggests a metric based on probabilistic theory to improve the interpretability of the values of the metric. In this context, currency can be interpreted as the *probability* that an attribute value is still up to date. Heinrich et al. [2007] assume that the shelf life of attribute values is exponentially distributed (random variable). The exponential distribution is a typical distribution for lifetime. However, this assumption also does not hold for all attributes (we will discuss this later). The proposed metric bases on two parameters *age*(*w, A*) and *decline*(*A*). The first parameter denotes the age of the attribute value *w*, which is derived by means of two factors: the instant of quantifying DQ and the instant of data acquisition. The average decline rate *decline*(*A*) of values of attribute *A* can be determined statistically. Heinrich et al. [2007] employ the metric on the level of attribute values as:

$$Currency = \exp(-decline(A) \cdot age(w, A))$$

Thus, after having determined the decline rate once, *currency* can be automatically quantified for each attribute value (R 6) using the metadata to determine the parameter *age*(*w, A*). In addition, the value for *currency* (as defined above) denotes the probability that the attribute value is still valid. This interpretability (R 3) is an advantage compared to the approaches mentioned above. Moreover, cases where *decline*(*A*) = 0 (for instance attributes like *date of birth* or *place of birth*, which never change) are taken into account correctly:

$$Currency = \exp(-decline(A) \cdot age(w, A)) = \exp(0 \cdot age(w, A)) = \exp(0) = 1$$

The same holds for $age(w, A) = 0$ (the attribute value is acquired at the instant of quantifying DQ):

$$Currency = \exp(-decline(A) \cdot age(w, A)) = \exp(-decline(A) \cdot 0) = \exp(0) = 1$$

Thereby the metric meets the requirements normalization (R 1) and interval scale (R 2). Moreover, formulas are provided to aggregate the values of the metric to higher levels (R 4). The metric is also adaptable to the context of a particular application as it allows incorporating weights to emphasize particular attributes and relations. Regarding this aspect of adaptivity, R 5 is met. Summarizing, the metric meets all the requirements stated above, if the shelf life of attribute values considered is exponentially distributed with the parameter *decline*(*A*).

However, this last assumption can be criticized, as the exponential distribution is memoryless in the following way:

$$P(X \geq x + t \mid X \geq x) = P(X \geq t)$$

If an exponentially distributed random variable $X$ exceeds the value $x$, then exceeding $x$ by at least $t$ is as probable as the exponentially distributed random variable $X$ exceeding the value $t$. In the context of DQ this means: The probability that a particular attribute value becomes out of date is equally high for each moment within a particular time period. Hence, this probability is independent from the current age of the attribute value. If two attribute values $a$ and $b$ are up to date at the instant of quantifying, then the probability of becoming out of date within the subsequent period of time is the same for both values, even if – for instance – $a$ is much older than $b$. It is obvious that assuming the shelf life of an attribute to be exponentially distributed does not hold for all attributes (for instance for different values of the attribute *professional status* within a customer database). Therefore, the metric in Heinrich et al. [2007] is not applicable within all contexts and R 5 is only partly met.

Table IV sums up the results regarding to the metrics discussed above.

Table IV. Evaluation of existing metrics for currency

| | **[Hinrichs 2002]** | **[Ballou et al. 1998]** | **[Even and Shankaranarayanan 2007]** | **[Heinrich et al. 2007]** |
|---|---|---|---|---|
| Technical definition of currency | $\frac{1}{(mean\ attribute\ update\ frequency)(age\ of\ attribute\ value)+1}$ | $\{\max[(1-\frac{age\ of\ attribute\ value}{shelf\ life}),0]\}^{s}$ | $1-\left(age\ of\ attribute\ value/T_{\max}\right)^{s}$ [example] | $\exp(-decline(A)\cdot age(w,A))$ |
| R 1. [Normalization] | Yes | Yes | Yes | Yes |
| R 2. [Interval scale] | No | Partly (for $s$ = 1) | Partly (for cardinal scaled utility functions) | Yes |
| R 3. [Interpretability] | No | Partly (for $s$ = 1) | Partly (as utility) | Yes |
| R 4. [Aggregation] | Yes | No | Yes | Yes |
| R 5. [Adaptivity] | Partly | Partly | Yes | Partly |
| R 6. [Feasibility] | Yes | No | Partly | Yes |

Due to the shortcomings of the discussed approaches, we now present a procedure to develop metrics for currency which takes into account the requirements to a larger extent.

## 3. DEVELOPING DATA QUALITY METRICS FOR CURRENCY

The existing metrics for currency either do not explicitly take into account the specific characteristics and thus the distribution of the shelf life of attribute values or assume a particular distribution. Therefore, they are not suitable for a number of important attributes. As most attributes differ in the specific characteristics of the shelf life (as defined by Ballou et al. [1998], cf. above) of their values, it is not possible to provide one technical definition which represents the metric for currency (as most of the existing approaches do – see Table IV). Instead, we design a procedure to develop metrics for currency which we consider as our artifact according to the guidelines for conducting design science research by Hevner et al. [2004]. The result of the procedure shall be an adequate metric which meets all six requirements. When designing this procedure, we particularly address adaptivity (R 5) and avoid the disadvantages of existing approaches concerning this requirement (cf. Table IV). The procedure follows the probabilistic approach presented in Heinrich et al. [2007] and takes into account the distribution of the shelf life of the attribute to be valuated. Thereby, the resulting metric can be adapted to the specific characteristics of the shelf life of the attribute in question. This allows elimination of limiting assumptions about the shelf life.

Our procedure consists of six steps (cf. Figure 2). In the following we describe these steps and how they are designed related to the requirements R 1 to R 6.

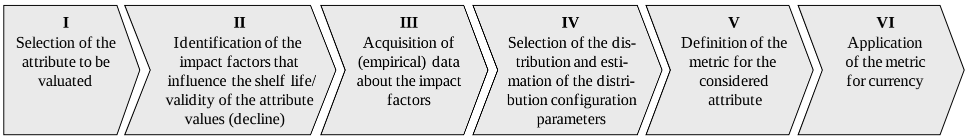


Fig. 2. Procedure to develop metrics for currency

In step I the attribute to be valuated is selected. Hereby – according to R 6 – we have to analyze from an economic point of view if the development of a metric for a particular attribute is profitable with respect to the given purpose (note that the development of a metric can be very costly, whereas the step of quantifying DQ itself may be automated). Hence, one should focus on attributes which are relevant for the task at hand. For instance, it is not necessary to develop metrics for *all* customer attributes within a customer relationship management (CRM) campaign. Instead, one should focus on the attributes to be used as selection criterion to identify customers for the target group of campaigns. Only if the given purpose justifies the quantification of DQ, step II should be conducted for a particular attribute.

Before acquiring data, factors which influence the shelf life of the attribute values (i.e. what does the decline rate of the attribute values depend on?) have to be determined in step II. The more impact factors are considered, the more exact the results of the metric will be (and vice versa). Thus, with respect to R 5, we can improve the metric by excluding or including particular impact factors. Taking into account R 6, this decision also depends on the costs for determining the impact factors.

If more than one impact factor is selected in step II, steps III to V have to be carried out for each factor. The following outline of the steps III and IV does not explicitly refer to the requirements R 1 to R 6. However, the described activities are needed to (1) assure a value of the metric which can be interpreted as a probability (R 1 to R 3) (2) adapt the metric to the specific characteristics of the shelf life of the attribute under consideration (R 5).

In step III data on the decline rate of the shelf life of the attribute values have to be acquired. Sources for such data might be external statistics (e.g. Federal Statistical Offices or scientific studies). Regarding the attribute *last name* empirical data from the Federal Statistical Office of Germany considering marriages/divorces may be taken into account, for example. If no such third party data are available, company-own (historical) data may be analyzed. For instance, historical customer data could be extracted from the data warehouse of the company to determine the decline rate of the attribute *current tariff* (i.e., how long does a customer use a certain tariff on average?). If neither external nor internal data on the decline rate of the shelf life of the attribute values are available, there are two possibilities: On the one hand, a sample of the customer base might be drawn. These customers could be surveyed in order to get data on the shelf life of the attribute values. These data can then be used to determine the average decline rate, which can be used for developing the metric. A short example considering the quality of address data may illustrate this: If information is needed to determine the frequency of relocation, it would be possible to draw a sample of persons (e.g. taken from the customer base) and survey them. After having determined the average duration of validity of a customer address (i.e., how long does a customer live in the same habitation on average?), the average decline rate can be calculated by means of the quotient (1/(average duration of validity of the addresses)). On the other hand, decline rates based on experts' estimations may be used. For instance, instead of using historical data, key account managers could also be surveyed considering the decline rate of the attribute *current tariff*.

In step IV adequate probability distributions for the shelf life of the attribute values have to be chosen based on the data acquired in step III. When choosing a distribution, one has to bear in mind the properties of this distribution. Table V states important properties of selected cumulative distribution functions:

Table V. Important properties of selected cumulative distribution functions

| Cumulative distribution function | Properties | Example |
|---|---|---|
| **Uniform distribution:**<br>A random variable being equally distributed over [a; b] $X$~U($a$; $b$) possesses the following cumulative distribution function:<br>$F(x)=\begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} \quad \text{for} & a < x < b \\ 1 & x \ge b \end{cases}$ | – Not memoryless<br>– Constant, absolute decline rate<br>– Fixed maximum period of validity and shelf life<br>– Continuous distribution | Analyses on the validity of customers' debit or eurocheque cards (each card has an unknown date of issue and a fixed expiry date) |
| **Exponential distribution:**<br>An exponentially distributed random variable $X$ with rate parameter $\lambda$ is characterized by the following cumulative distribution function:<br>$F(x)=\begin{cases} 1-\exp(-\lambda x) \quad \text{for} & x \ge 0 \\ 0 & x < 0 \end{cases}$ | – Memoryless (i.e. the conditional probability that the attribute value becomes out of date in the next period of time is independent of its current age)<br>– Constant, relative decline rate<br>– Continuous distribution | Analyses on currency of address data (e.g. relocation) |

| | | |
|---|---|---|
| **Geometric distribution:**<br>A geometric distributed random variable $X_n$ with parameter $q = 1 - p$ (with $q$ as probability for a failure) possesses the following cumulative distribution function:<br>$F(n) = 1 - (1-p)^n$ | – Memoryless<br>– Constant, absolute probability of decline within each period<br>– Discrete distribution | Analyses on the validity of a contract (e.g. labor agreement) with the option to terminate at quarter-end |
| **Weibull distribution:**<br>A Weibull distributed random variable $X$ with shape parameter $k > 0$ and scale parameter $\lambda > 0$ has the following cumulative distribution function:<br>$F(x) = \begin{cases} 1-\exp(-\lambda x^k) & \text{for } x \geq 0 \\ 0 & x < 0 \end{cases}$ | – Not memoryless<br>– Applicable for increasing, decreasing and constant relative decline rates<br>– Continuous distribution | Analyses on the duration of study and professional status *student* (cf. examples below) |
| **Gamma distribution**<br>A gamma distributed random variable $X$ with shape parameter $k > 0$ and scale parameter $\theta > 0$ has the following cumulative distribution function:<br>$F(x) = \begin{cases} \dfrac{\gamma\left(k, \frac{x}{\theta}\right)}{\gamma(k)} & \text{for } x \geq 0 \\ 0 & x < 0 \end{cases}$ | – Not memoryless<br>– Applicable for the description of changing, relative decline rates<br>– Continuous distribution | Analyses on the lifespan of end devices (e.g. within marketing campaigns for accessories) |

Distributions can usually be adapted via distribution configuration parameters (for instance, the shape and scale parameter of the Weibull distribution mentioned in Table V). These parameters determine the decline rate and have to be calculated by means of common estimation procedures applied to the data (supported by statistical analysis software like e.g. R or SPSS)[2].

In cases where several factors have an impact on the decline rate, it is not sufficient to conduct steps III-IV for each factor. Moreover, the distribution functions have to be combined. This is done in step V, in which the metric is defined based on the combined distribution (this ensures that the requirements R 1, R 2 and R 3 are fulfilled, as the result is a probability). The developed metric quantifies currency on the level of attribute values. To allow a flexible application and enable a quantification of DQ at the level of tuples, relations (especially views), and the whole database (cf. R 4), the metric can be aggregated to the next higher levels as shown in Heinrich et al. [2007]. R 5 is met due to the following reasons. The first one concerns the level of attribute values: The procedure takes into account the distribution of the shelf life of the attribute to be valuated. Hence, the procedure is designed in such way that the resulting metric can be adapted to the characteristics of each particular attribute (e.g. constant, increasing, decreasing, and changing relative decline rates of its shelf life). This is necessary to enable an objective

[2] For determining adequate probability distributions, we refer the interested reader to existing literature on kernel density estimation (e.g. [Parzen 1962; Wasserman 2004]) and confidence intervals (e.g. [O'Gorman 2004]).

measurement for all attribute values considered in a certain business context. The second reason concerns the next higher levels: When aggregating the values of the metric on the level of attribute values to the next higher level (tuples), certain attributes can be emphasized as described in Heinrich et al. [2007] to adopt the measurement to a particular business context. Consider the following example: During two different campaigns, the customer shall be contacted either by phone (campaign 1) or by mail (campaign 2). When quantifying currency of a customer's data, i.e., on the level of tuples, obviously the attributes "customer's phone number" and "customer's postal address" differ in their importance for the two campaigns. For the phone campaign (1), the customer's postal address is irrelevant. Therefore, "customer's postal address" shall be assigned a weight of 0. In turn, for the postal campaign (2), the customer's phone number is not needed. Hence, "customer's phone number" has to be weighted with 0. This weighting is utterly necessary, otherwise the metric is not applicable or we get unreasonable results.

In Step VI currency is quantified by means of the developed metric. As this step shall be accomplished at the highest possible level of automation (cf. R 6), the age of each attribute value has to be determined automatically. Therefore, the age is calculated (e.g. using SQL DML-statements) as the time period between the instant when DQ is quantified and the instant of data acquisition, as both pieces of information are usually stored as metadata in a database. Afterwards, the value of the metric for currency is calculated using the combined distribution function of step V. Finally, the results of the metric can be applied in an economic management of DQ (see next sections).

Table VI summarizes the individual steps of the procedure and denotes the corresponding tasks.

Table VI. Procedure to develop metrics for currency and important tasks

| Step | | Tasks |
|---|---|---|
| I | Selection of the attribute to be valuated | - Analyze for which attributes the quantification of currency is necessary or useful related to the defined goal.<br>- Identify those attributes, for which developing a metric for currency makes sense from a cost/benefit point of view (*economic* management of DQ). |
| II | Identification of the impact factors that influence the shelf life/validity of the attribute values | - Identify the impact factors which influence the decline rate of the attribute values.<br>- Decide in a goal-oriented way which impact factors shall be taken into account when developing the metric for the attributes identified in step I. |
| III | Acquisition of data on the impact factors | - Access data on the impact factors to analyze the decline rate for values of the attribute chosen in step I. Such data are e.g.:<br>(1) publicly accessible data (e.g. Federal Statistical Offices)<br>(2) already existing company-own data (e.g. data warehouse)<br>(3) survey/interview company-own customers to get the necessary data (e.g. shelf life of the attribute values)<br>(4) data, especially decline rates, based on experts' estimations |

| | | |
|---|---|---|
| IV | Selection of the distribution and estimation of the distribution parameters | - Determine suitable probability distributions for the shelf life of attribute values based on the data accessed in step III.<br>- Estimate the corresponding distribution parameters taking into consideration the impact factors identified in step II. |
| V | Definition of the metric for the considered attribute | - Combine the distribution functions from step IV.<br>- Define the metric, which returns the probability that the values of the attribute are still up to date. |
| VI | Application of the metric for currency | - Quantify DQ using the metric in an automated way (by means of SQL statements etc.).<br>- Establish the metric to quantify currency as an important part of an *economic* DQ management within the company. |

We study the procedure in depth in business in section 5 and illustrate its application in a real world scenario. Before doing so, we briefly compare the development of the metric and the quantification of DQ to other metrics with respect to costs.

Requirement R 6 says: „When defining metrics, methods to determine the input parameters shall be defined. If exact determination of input parameters is not possible or too cost-intensive, alternative rigorous methods (e.g. statistical) shall be proposed. From an economic point of view, it is also required that the quantification of DQ can be accomplished at a high level of automation.“ Referring to currency, exact determination means to compare the considered attribute values to their real world counterparts and to verify, whether the values stored in the information system are still valid (related to the DQ dimension correctness). It is obvious that – for large data sets (for instance, 156,000 stored customer data sets in the real world example in section 5) – this verification is much too expensive and time consuming. Hence, according to R 6, we have to look for alternative methods. Hinrichs [2002], Ballou et al. [1998], Even and Shankaranarayanan [2007] and Heinrich et al. [2007] developed corresponding approaches. However, they are based on assumptions which impede an adequate quantification of currency for many attributes (cf. section 2). Even if these procedures could be accomplished at lower costs than the approach designed in this paper, they would lead to unreasonable or actually wrong results.

That is why we proposed a procedure to develop metrics which allows the step of quantifying DQ to be automated (application of the metric). It can be costly to develop such a metric. However, it is much less time- and labor-intensive than checking for each and every customer, whether his/her data is still up to date. Instead of this we can use public information for a number of attributes (e.g. from Federal Statistical Offices or scientific studies). Regarding the attributes *marital status* and *address,* empirical data from the Federal Statistical Office considering marriages/divorces and the frequency of relocation can be taken into account, for instance. If such publicly available data have been acquired, the metric can be developed and applied to all values of an attribute several times. Whereas a manual verification causes very high costs every time it is done, the costs of data acquisition are allocated to several uses. It also may seem costly to identify the distribution and to estimate the distribution configuration parameters. However, statistical analysis software (like e.g. R or SPSS) reduces the corresponding costs, as this kind of software proposes distributions and distribution configuration parameters based on the empirical data. To put it in a nutshell: compared to other procedures (which lead to unreasonable results for many attributes) the designed procedure is by far less time- and labor-intensive.

## 4. ILLUSTRATION OF THE PROCEDURE

In this section we illustrate the procedure by means of an example and develop a particular metric for the DQ dimension currency. As a continuing example we use the attribute *professional status* within a customer database (step I). We chose this attribute as it is also part of our real world example in cooperation with a major mobile services provider (MSP) in the next section where we consider a CRM campaign in which students shall be addressed. The problem of such campaigns is that customers are often included in the target group, as their professional status is stored as *student* in the database, although they have already finished or abandoned their studies. In this case they cannot be granted a student discount anymore. The implications of selecting wrong customers for the target group are twofold: decreased customer satisfaction on the one hand and low success rates of campaigns on the other hand lead to inefficient usage of resources. To reduce such problems, we present a metric for the attribute *professional status* with the attribute value *student*, which is a selection criterion for the target group of the campaign (step I, goal-oriented selection of the attribute to be valuated).

The attribute value *student* can lose its validity due to two impact factors (cf. step II): A study is either completed or aborted. As neither of them can be neglected, the metric consists of two different "characteristics" and distributions, one for each impact factor. For the problem at hand we need neither a sampling survey nor any other form of internal data collection. Instead, the distributions can be determined by means of external data: Many universities as well as Federal Statistical Offices provide statistics on the duration of studies (step III, cf. [Hackl and Sedlacek 2001; Federal Statistical Office of Germany 2006]). For illustrational purposes we use data from the University of Vienna (Austria): They provide a relative frequency distribution of the duration of study, which aggregates the data of several courses of study for different faculties.

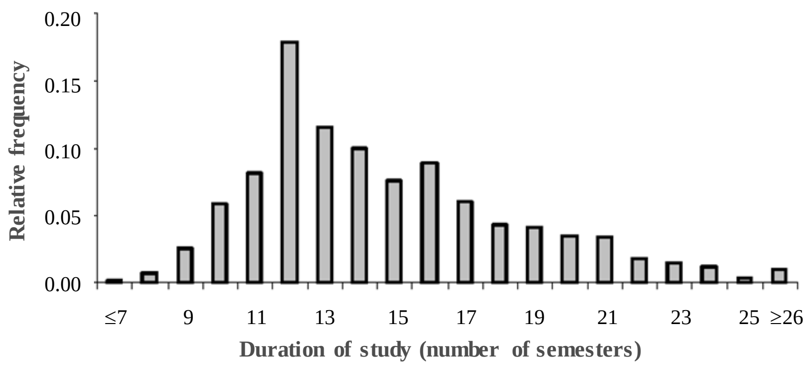


Fig. 3. Relative frequency distribution of duration of study [Hackl and Sedlacek 2001]

Figure 3 depicts the relative frequency distribution concerning the duration of study at the University of Vienna (for all students graduating in 2000). Considering the first impact factor (successful completion of degree), we can determine the distribution of the duration of study. In this case, assuming a constant relative decline rate would imply the following: The probability that a student who has already been studying for eight

semesters will complete his degree within the next two semesters is equal to the probability that a student who has already been studying for twelve semesters will complete his degree within the next two semesters. This obviously does not hold as initially the relative frequency steeply increases and decreases again after the 12th semester (cf. figure 3). That is why we can assume neither a constant, relative decline rate nor memorylessness – both important properties of the exponential distribution (cf. Table V). Hence, the approaches by Hinrichs [2002], Ballou et al. [1998], Even and Shankaranarayanan [2007] and Heinrich et al. [2007] are not suitable within this context.
Therefore we need a distribution of the shelf life which is not memoryless and does not assume constant decline rates (step IV). A continuous distribution holding these properties is the Weibull distribution, for instance (cf. Table V).

The Weibull distribution *wei*($k$, $\lambda$) is based on two parameters, *shape* ($k$) and *scale* ($\lambda$). A number of alternatives exist to determine these parameters for the problem at hand. Marks [2005] presents a method to determine the Weibull distribution parameters based on symmetric percentiles $P_L$ and $P_U$ (lower and upper percentile $L$ and $U$, denoting the value of the distribution at the percentile $P_L$ and $P_U$ respectively). Percentiles are the values of a variable below which a certain percentage of observations fall. An example for symmetric percentiles is the 10th percentile ($P_{10}$) and the 90th percentile ($P_{90}$). For this example 90% of all values lie below the 90th percentile. The simple estimation is based on the following equations for $k$ and $\lambda$ (with *ln* as natural logarithm function):

$$k = \ln\left(\frac{\ln(U)}{\ln(L)}\right) \cdot (P_U / P_L) \text{ and } \lambda = \frac{P_U}{(-\ln(L))^{1/k}}$$

By applying a Monte Carlo simulation, Marks illustrates that the best estimation is achieved when using the 10th and 90th percentile. We can utilize this method, but have to adapt it slightly: Marks implicitly assumes that the Weibull distributed values start at the point of origin. However, in our example the graduates complete their degrees between the 7th and 26th semester. That is why the calculated parameters have to be adjusted by means of a left shift. This way we get $k = 0.00002$ and $\lambda = 4$ based on the data presented by Hackl and Sedlacek [2001]. The coefficient of determination $R^2$ is 0.91, expressing that the parameterized Weibull distribution approximates the empirical distribution adequately. As a consequence, the cumulative distribution function can be formulated as follows:

$$P_{Gradute}(x) = 1 - \exp(-0.00002 \cdot x^4) \text{ for } x \geq 0$$

$P_{Graduate}(x)$ denotes the cumulative probability that a student has completed his degree after $x$ semesters (step IV).

Furthermore, we have to analyze the distribution of dropouts as the second impact factor on the validity of the attribute value *student.* Figure 4 illustrates the corresponding data for the University of Vienna (step III). It shows the percentage of all dropouts that aborted their studies within a particular semester (again aggregated for all programs of study): For instance, about 18% of all dropouts discarded their studies in the first semester. It holds for this distribution that the dropouts' percentage remains approximately constant in relation to the students still active (in contrast to the number of absolute dropouts, which is obviously decreasing). Hence, we state approximate memorylessness as well as a constant relative decline rate. Therefore, we can apply the exponential distribution.

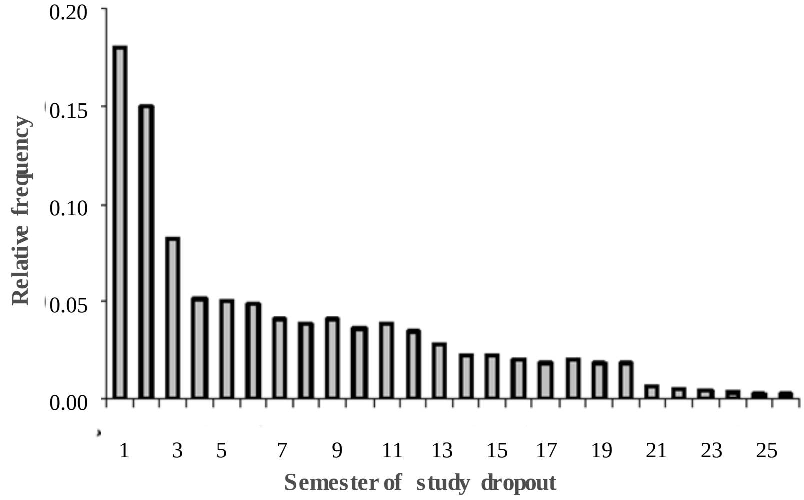


Fig. 4. Relative frequency distribution of study dropout [Hackl and Sedlacek 2001]

To estimate the parameters (step IV) for the exponential distribution, we can make use of the expected value: It corresponds to the reciprocal of the decline rate. The arithmetic mean of the empirical data serves as unbiased estimator for the expected value $E(x)$. For the given data, the arithmetic mean is about 5.5 semesters. Thereby, the distribution parameter $\lambda$ of the exponential distribution is calculated as follows:

$$E(x) = \frac{1}{\lambda} \Leftrightarrow \lambda = \frac{1}{E(x)} \approx \frac{1}{5.5} = 0.18$$

Again, we get an adequate approximation of the empirical distribution by means of the parameterized exponential distribution ($R^2 = 0.88$). $P_{Dropout}(x)$ denotes the cumulative probability that a student aborted his study (step IV).

$$P_{Dropout}(x) = 1 - \exp(-0.18 \cdot x) \text{ for } x \geq 0$$

To integrate the two distributions determined above we have to estimate the percentage of graduates and dropouts. Using empirical data, the percentage of graduates can be estimated at 64%. Hence, the probability $P_{Student}(x)$ that a student is still studying can be defined as follows (step V):

$$P_{Student}(x) = 1 - 0.64 \cdot (P_{Gradute}(x)) - 0.36 \cdot (P_{Dropout}(x)) \text{ for } x \geq 0$$

We can use this formula as a metric for currency and calculate the probability that a customer with the professional status *student* (within a database) is still studying. Based on probability theory, the values of the metric are normalized to [0; 1] (R 1), interval scaled (R 2) and interpretable (R 3). Moreover, the aggregation formulas defined by

Heinrich et al. [2007] can be applied (R 4). The weighting factors in these aggregation formulas and designing the metric to the shelf life of the attribute value (and the impact factors) make the metric meet R 5. As mentioned earlier, the currency of a particular customer's professional status can be calculated automatically (R 6) by using the formula above.

After illustrating the individual steps of the procedure to develop a metric for currency, the application of the metric within the mobile services sector is described in the next section.

## 5. REAL WORLD EXAMPLE: APPLYING THE METRIC FOR CURRENCY

This section illustrates the application of the metric for currency "in depth in business" [Hevner et al. 2004, p. 86]. Our objective is to point out the economic effects of quantifying currency with our metric within CRM campaigns by means of a detailed real world example. Particularly, we want to illustrate that using the values of the metric when selecting the target group of customer campaigns leads to higher profits.

The metrics for currency were developed according to the procedure designed above. We exemplify the procedure by means of a particular attribute and its characteristics, but the results are reusable: If the attribute shall be used for other tasks (e.g. for designing new products and tariffs), the metric needs not to be developed again.

In our real world example we focus on a specific CRM campaign of the mobile services provider (MSP) that faced the following problem: The provider wanted to offer a new premium tariff called *Student AAA* to customers with the professional status *student*. The new tariff increases the MSP's return on sales by 5%. For reasons of confidentiality, all specific figures and data had to be changed and made anonymous. Nevertheless, the procedure and the basic results remain the same.

Before the DQ project, the target group of such campaigns was selected from the customer database as follows:

1. Select all customers who fulfill a given selection criterion (e.g. attribute value for professional status *student*).
2. Rank the selected customers according to their sales volumes.
3. Select the top X% customers out of the ranked customers, which constitute the target group of the campaign.

After that, the new offer was sent to the customers of the target group. In the past, the a posteriori success rates of such campaigns averaged out at approx. 9%. This means, about 9 out of 100 addressed customers accepted the new offer.

Applying this *previous* selection procedure to the new campaign *Student AAA* means to select the top 30% customers with regard to their sales volumes out of all customers with the professional status *student* (note that 30% was a requirement from the marketing department). Thus, about 46,800 customers (out of all 156,000 customers with the attribute value *student* as professional status) would be addressed. These 46,800 customers possess an average sales volume of 1,340 € p. a. Assuming the former success rate of about 9%, the number of customers who will accept the offer can be estimated at about 4,200. 4,200 customers with average sales volumes of 1,340 € accepting the offer would imply a forecasted additional profit of approx. 281,400 € (=4,200*1,340 €*5%) which sounds like a quite profitable business case for the campaign.

Yet before starting the campaign and addressing these 46,800 customers, its profit should be improved by means of the DQ metric developed above. Especially the success rate should be increased by raising the percentage of addressed customers who are indeed still studying (only they can accept the offer as one needs to provide a certificate of matriculation).

To achieve this, the value of the metric was calculated for each of the selected 46,800 customers (top 30% according to their sales volumes) with the professional status *student.* These metric values indicate the probabilities that the corresponding customers are still studying in reality. This step had to be done automatically, since a manual "check" of each of the 46,800 customers would have been – as described earlier – far too laborious and cost-intensive.

The value range of the metric [0; 1] was divided into ten intervals ([0; 0.1[, [0.1; 0.2[, …, [0.9; 1.0]) and the selected 46,800 customers were assigned to these intervals according to their individual value of the metric. Figure 5 depicts the number of customers within each interval. For example, the value of the metric in the interval [0.2; 0.3[ was for approx. 3,500 customers.

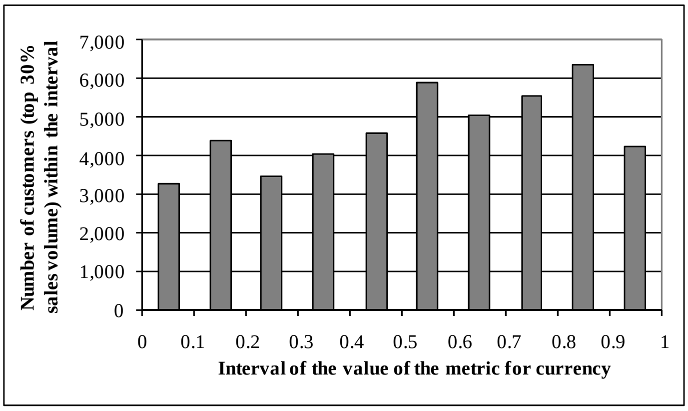


Fig. 5. Numbers of customers depending on their value of the metric for currency

As can be seen from the figure, the probability that the selected customers are still studying in reality (value of the metric) is less than 0.5 for approx. 20,000 customers (about 42%). Because of that, it is very unlikely that these customers are able to accept the offer at all.

To overcome this problem, a *new* selection procedure based on the developed metric for currency was defined: Firstly, the value of the metric was automatically calculated for all 156,000 customers with the professional status *student*. Then the expected additional profit was computed for each customer based on

1. his/her individual sales volume p.a.,
2. his/her individual value of the metric for currency (probability that he/she is still studying in reality), and
3. the possible additional return on sales of 5% *in case the offer is accepted*.

After that, this expected additional profit was used as a selection criterion to identify the top 30% customers. Obviously, this new selection procedure chose other customers than the previous selection procedure, as now the currency of the attribute values was taken into account: Some of the customers with the professional status *student* showing high sales volumes were very likely not to have student-status in reality anymore. Using the expected additional profit as selection criterion, these particular customers were not selected anymore. In the end, only approx. 18,100 customers were selected according to both selection procedures (sales volume and expected additional profit). This means: More than 28,700 customers, who would have been addressed based on the previous selection procedure of the MSP, were not addressed anymore when taking the metric for currency into consideration. Instead, 28,700 other customers were identified for the campaign based on the criterion expected additional profit.

As a precaution, the marketing department of the MSP decided to address all approx. 75,500 customers who were selected according to one or both of the selection procedures. This was done in order to verify whether including DQ provides better results. The analysis conducted after the campaign revealed the following results (cf. Table VII) for the previous and the new selection procedure (based on the DQ metric for currency):

Table VII. Results of the real world example

| | **Previous selection procedure** | **New selection procedure (based on DQ metric)** |
|---|---|---|
| Selection process for the target group of the campaign | (1) Select all customers whose professional status is stored as *student*<br>(2) Rank the selected customers according to their **sales volumes**<br>(3) Select the top 30% out of the ranked customers for the target group | (1) Select all customers whose professional status is stored as *student*<br>(2) Rank the selected customers according to their individual **expected additional profits** based on (a) individual sales volume (b) individual value of the metric for currency (c) additional return on sales (5%) in case the offer is accepted<br>(3) Select the top 30% out of the ranked customers for the target group |
| Number of Customers with attribute value *student* as professional status | 156,000 | 156,000 |
| Number of Customers selected for the campaign | 46,800 | 46,800 |
| Number of Customers accepting the offer | Approx. 2,660 | Approx. 5,380 |
| Success rate | Approx. 5.7% | Approx. 11.5% |
| ∅ Sales volumes per customer | Approx. 1,340 €[3] | Approx. 1,150 € |
| Additional profit of the campaign | Approx. 178,200 € | Approx. 309,200 € |

The table illustrates that the overall success rate of the 46,800 customers who were initially selected by sales volume was a posteriori only about 5.7% and therefore even lower than expected. This can be explained as follows: Indeed, according to the previous procedure the customers with the highest sales volumes were selected. However, a

[3] For reasons of simplicity, the ∅ a priori sales volume per customer were also used for the a posteriori calculation.

posteriori many of these customers were not studying anymore and therefore could not accept the new offer. From an economic point of view it is highly questionable to contact these customers when taking into account the customer-contact costs. Considering the customers selected in the new procedure using the expected additional profit as selection criterion, we can state the following: On the one hand, the average sales volume per customer (1,150 €) was lower (previous procedure: 1,340 €). On the other hand, the success rate of about 11.5% exceeded the expectations and was much higher than with the previous procedure (5.7%). In the end, the MSP made an additional profit of approx. 309,200 €. In contrast, by addressing only the customers with the highest sales volumes, the additional profit would have been far lower (178,200 €).

## 6. SUMMARY

In subsection 6.1 we discuss the findings of the real world example in section 5. A summary of the contributions of the paper, the limitations, and possible directions for future research follow in subsection 6.2.

### 6.1 Findings of the real world example

The real world example illustrates the applicability of the metric for currency and gives an example how the metric can help to improve to substantiate the calculation of business cases and thereby the results of CRM campaigns. It becomes clear that using only sales volumes as selection criterion results in choosing many customers who are very likely not to study anymore. Hence, many of the addressed customers are not eligible to accept the offer. Such campaigns are less profitable than campaigns which take into account the currency of the customer data. By applying the metrics, the MSP was not only able to improve the success rates of the campaign. In addition, it could establish a correlation between the success rates of campaigns and the results of quantifying DQ.

The decisive advantage of the designed procedure is its foundation in probabilistic theory. If we interpret currency as the probability that an attribute value still corresponds to the current state of its real world counterpart, this probability (for the occurrence of an event; for instance, professional status *student* is up to date or not) can be used in decision situations. The MSP was able to calculate an expected additional profit, leading to a substantiated and comprehensible decision support. Moreover, the process for selecting customers was improved significantly for the campaigns of the MSP, a fact that helped to cut down campaign costs. Finally, the MSP will be able to estimate the economic benefit of DQ measures in a more accurate way in the future.

The following table summarizes the evaluation steps we used in this paper:

Table VIII. Evaluation step and description

| **Evaluation step** | **Description** |
|---|---|
| Requirements to design and evaluate the artifact | - Definition of requirements for DQ metrics derived from literature<br>- Analysis whether the metrics developed according to the designed procedure meet the requirements<br>(Both steps conducted before the real world example) |
| Integration of the artifact within business environment | - Application of the metric developed according to the designed procedure in a mailing campaign of a mobile service company |

| | |
|---|---|
| Definition of appropriate metrics | - Comparison of the economic results of two procedures for selecting the target group of a mailing campaign<br>- Key figures of the new procedure: (1) Higher success rate of the campaign (2) Additional profit of the campaign |

One could criticize that we base our findings on a single project. Hence, the application of the procedure should be repeated (especially in other areas) to substantiate our findings. Moreover, we need to emphasize that the campaign only considered those customers whose professional status was already stored as *student* in the database. However, there are also customers who have become students (in the meantime since their professional status was acquired), but whose attribute value in the database is different (e.g. *pupil, apprentice*). Thus, they may be candidates for the target group of the campaign as well. Since these customers are not considered yet, the existing metric for currency has to be extended/adapted for other attribute values (as for instance *pupil*) which indicate the probability for a transition into the professional status *student*. Again, data from the Federal Statistical Offices can be used.

### 6.2 General results, limitations and further research

The paper analyzed how the DQ dimension currency can be quantified in a goal-oriented and economic manner. The aim was to design a procedure to develop metrics for currency. The metrics enable an objective and partly automated quantification. In contrast to existing approaches, the metrics developed according to the procedure meet important requirements like interpretability and feasibility. Moreover, they are not based on (limiting) assumptions as for instance an exponential distribution for the shelf life of the attribute values considered. They enable quantifying DQ and can represent the basis for economic analyses. The effect of DQ measures can be analyzed by comparing the realized DQ level (a posteriori) with the planned level (a priori). Figure 6 depicts the DQ loop. It depicts the process of comparing both a posteriori and a priori level as well as the resulting economic effects. It illustrates that *DQ measures* (like data cleansing measures, buying external address data etc.) are applied to *data values* to improve the *realized DQ level*. The relevant question is: To what extent should such measures be taken and applied to data values whose validity *declines* (disturbance variable) over time? To put it another way: To which extent is taking such measures justified from an economic point of view?

To support such decisions, the current level of DQ has to be quantified as DQ measures are often most useful for attribute values of low quality (cf. the real world example in section 5). Basically, there are two options of quantifying DQ: On the one hand, the quality of each attribute value can be verified by means of a *real world test*. This corresponds to quantifying the DQ dimension *correctness*. The other option is an *estimation* of the quality of the attribute values. This corresponds to quantifying *currency* (as defined earlier). Both ways cause different amounts of *costs*. Especially in case of large data sets, a real world test is not practicable or not possible at all (cf. above). Therefore, estimating DQ by means of a metric for currency is normally the cheaper and better option to determine both, the current DQ level in a first step and the *planned DQ level* in a second step considering cost/benefit aspects (for instance, which customers should be addressed during a mailing campaign, which ones should not?). Based on the planned level of DQ, we can fix the extent of *DQ measures* to be taken (controller in the DQ loop). The balance between (estimated) *costs* and *benefits* per option can be used in two ways: We cannot only determine a level of DQ which is reasonable from an economic point of view, but also choose between the two options of quantification (as they essentially also influence the cost of DQ).

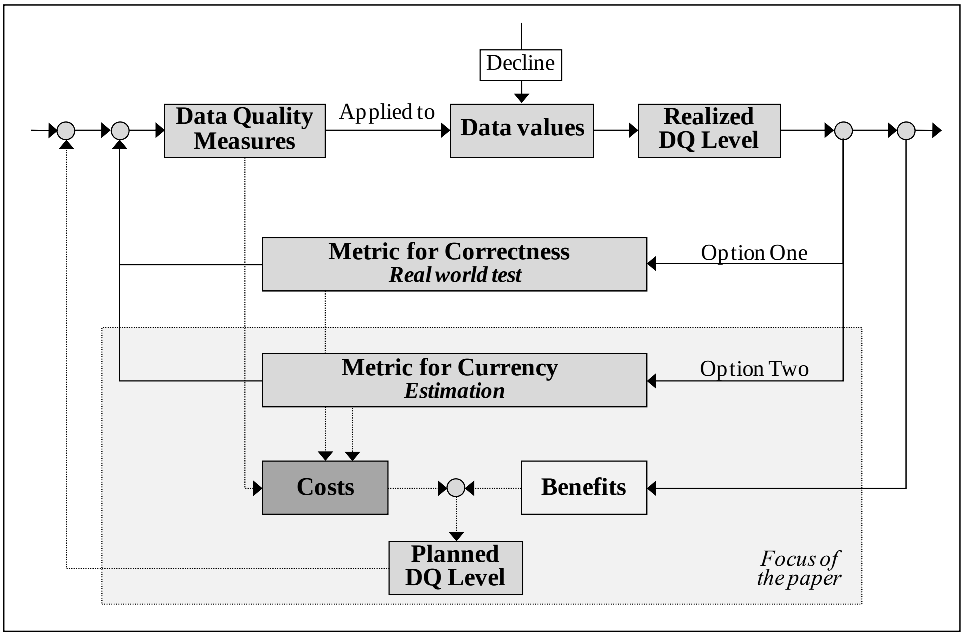


Fig. 6. Economics-driven data quality loop

Despite these improvements, some limitations of our current approach provide room for further research. One important prerequisite for developing metrics is a suitable data pool. In many cases, external data – for instance from Federal Statistical Offices or data provided by experts' estimations and forecasts – can help populate data pools. Conducting samples or other forms of internal data analysis is often more laborious and cost-intensive. Therefore it has to be examined whether the development of a metric for an attribute is necessary and reasonable with respect to the given goal. However, it has to be considered that a metric which was developed once can be reused several times or adapted to other fields of application. The authors currently work on a model-based approach for the economic planning of DQ measures. For implementing such a model, adequate DQ metrics and quantification procedures are necessary. The approach presented in this paper provides a basis for those purposes. Nevertheless, further metrics for other DQ dimensions should be developed and thus further research in this area is encouraged.

## REFERENCES

BALLOU, D.P., AND PAZER, H.L. 1985. Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems. *Management Science* 31(2), 150-162.

BALLOU, D.P., AND PAZER, H.L. 1995. Designing information systems to optimize the accuracy-timeliness tradeoff. *Information Systems Research* 6(1), 51-72.

BALLOU, D.P., WANG, R.Y., PAZER, H.L., AND TAYI, G.K. 1998. Modeling Information Manufacturing Systems to Determine Information Product Quality. *Management Science* 44(4), 462-484.

BATINI, C., AND SCANNAPIECO, M. 2006. *Data Quality. Concepts, Methodologies and Techniques (Data-Centric Systems and Applications).* Springer, Berlin.

CAMPANELLA, J. 1999. *Principles of Quality Costs: Principles, Implementation and Use.* ASQ Quality Press, Milwaukee.

CAPPIELLO, C., FRANCALANCI, C., AND PERNICI, B. 2004. Data quality assessment from the user's perspective. In *Proceedings of the IQIS 2004, International Workshop on Information Quality in Information Systems*, 68-73.

CAPPIELLO, C., FRANCALANCI, C., PERNICI, B., PLEBANI, P., AND SCANNAPIECO, M. 2003. Data Quality Assurance in Cooperative Information Systems: a Multi-dimension Quality Certificate. In *Proceedings of the ICDT'03 International Workshop on Data Quality in Cooperative Information Systems (DQCIS'03)*.

ENGLISH, L.P. 1999. *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits.* Wiley, New York.

EPPLER, M.J. 2003. *Managing information quality.* Springer, Berlin.

EVEN, A., AND SHANKARANARAYANAN, G. 2007. Utility-Driven Assessment of Data Quality. *The DATA BASE for Advances in Information Systems* 38(2), 75-93.

FEDERAL STATISTICAL OFFICE OF GERMANY 2006. *Education in Germany - Statistics of Higher Education.* Federal Statistical Office of Germany, Wiesbaden.

FEIGENBAUM, A.V. 1991. *Total Quality Control.* McGraw-Hill, New York.

HACKL, P., AND SEDLACEK, G. 2001. Analyse der Studiendauer am Beispiel der Wirtschaftsuniversität Wien. In Dutter, R. (ed.) *Festschrift 50 Jahre Österreichische Statistische Gesellschaft,* Österreichische Statistische Gesellschaft, Wien, 41-59. (In German)

HARRIS INTERACTIVE 2006. Information Workers Beware: Your Business Data Can't Be Trusted. Accessed 2008/12/18. http://www.businessobjects.com/news/press/press2006/20060626_data_quality_survey_comp.asp

HEINRICH, B., KAISER, M., AND KLIER, M. 2007. How to measure data quality? – a metric based approach. In *Proceedings of the 28th International Conference on Information Systems (ICIS).*

HEINRICH, B., KAISER, M., AND KLIER, M. 2008. Does the EU Insurance Mediation Directive help to improve Data Quality? – A metric-based analysis. In *Proceedings of the 16th European Conference on Information Systems (ECIS)*.

HELFERT, M., AND HEINRICH, B. 2003. Analyzing Data Quality Investments in CRM – A model-based approach. In *Proceedings of the 8th International Conference on Information Quality (ICIQ)*, 80-95.

HEVNER, A.R., MARCH, S.T., PARK, J., AND RAM, S. 2004. Design Science in Information Systems Research. *MIS Quarterly* 28(1), 75-105.

HINRICHS, H. 2002. *Datenqualitätsmanagement in Data Warehouse-Systemen.* Dissertation Universität Oldenburg, Oldenburg. (In German)

JARKE, M., AND VASSILIOU, Y. 1997. Foundations of Data Warehouse Quality – A Review of the DWQ Project. In *Proceedings of the 2nd International Conference on Information Quality (ICIQ)*, 299-313.

JURAN, J.M. 1998. How to think about Quality. In Juran, J.M., Godfrey, A.B. (eds) *Juran's Quality Handbook,* McGraw-Hill, New York, 2.1-2.18.

KLEIN, B.D., AND CALLAHAN, T.J. 2007. A comparison of information technology professionals' and data consumers' perceptions of the importance of the dimensions of information quality. *International Journal of Information Quality* 1(4), 392-411.

LEE, Y.W., STRONG, D.M., KAHN, B.K., AND WANG, R.Y. 2002. AIMQ: a methodology for information quality assessment. *Information & Management* 40(2), 133-146.

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

MARKS, N. 2005. Estimation of Weibull parameters from common percentiles. *Journal of Applied Statistics* 32(1), 17-24.

META GROUP 1999. *Data Warehouse Scorecard.* Meta Group

O'GORMAN, T.W. 2004. *Applied Adaptive Statistical Methods: Tests of Significance and Confidence Intervals.* Cambridge University Press, Philadelphia.

PARZEN, E. 1962. On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics* 33(3), 1065-1076.

PIPINO, L., LEE, Y.W., AND WANG, R.Y. 2002. Data Quality Assessment. *Communications of the ACM* 45(4), 211-218.

PRICEWATERHOUSECOOPERS 2004. Data Quality 2004 Survey. Accessed 2009/03/08. http://www.pwc.com/extweb/ncpressrelease.nsf/docid/A0C678E7D21B7EB3CA256F490008D3E5

REDMAN, T.C. 1996. *Data Quality for the Information Age.* Artech House, Boston.

REDMAN, T.C. 1998. The impact of poor data quality on the typical enterprise. *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.

SHANK, J.K., AND GOVINDARAJAN, V. 1994. Measuring the cost of quality: A strategic cost management perspective. *Journal of Cost Management* 8(2), 5-17.

STRONG, D.M., LEE, Y.W., AND WANG, R.Y. 1997. Data quality in context. *Communications of the ACM* 40(5), 103-110.

TEBOUL, J. 1991. *Managing Quality Dynamics.* Prentice Hall, New York.

WANG, R.Y., STOREY, V.C., AND FIRTH, C.P. 1995. A Framework for analysis of data quality research. *IEEE Transactions on Knowledge and Data Engineering* 7(4), 623-640.

WANG, R.Y., AND STRONG, D.M. 1996. Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems* 12(4), 5-33.

WASSERMAN, L. 2004. *All of Statistics: A Concise Course in Statistical Inference.* Springer, Berlin.

YU, Z., AND WANG, Y. 2007. An empirical research on non-technical factors related to statistical data quality in China's enterprises. *International Journal of Information Quality* 1(2), 193-208.