

ISO/IEC 15504 based assessments for Information Quality

Dirk Malzahn

OrgaTech GmbH, Zum Pier 73 – 75, 44536 Luene, Germany

dm@orgatech.org, phone +49 (0) 231 39995-305, fax +49 (0) 231 39995 444 305

Abstract

Discussing the topic of information quality leads to several problems. There are no common definitions of the terms “information” and “information quality” and a measurement framework for the quality of information does not exist.

On the other hand, most industries have a strong focus on information quality. Especially for the banking and insurance sector information quality is an indispensable element of their business quality.

In 2007 the German Society for Information and Data Quality published 15 dimensions of information quality. These dimensions are based on the studies of the MIT over the last 10 - 15 years and deliver the first complete definition of information quality.

However this definition can only be seen as a first step. The quality of information is not only determined by this definition but by the capability of an organization to handle this definition.

Here ISO/IEC 15504 can help. Even though the dimensions are no typical processes, an ISO/IEC 15504 based assessment can measure if the dimensions of information quality are sufficiently institutionalized, maintained and developed within the organization.

This article explains the dimensions of information quality and shows, how ISO/IEC 15504 can be applied to information quality, to reach a reliable and capable information quality organization.

1. Introduction

The quality of information is an inevitable element for most industries. Especially the banking and

insurance sector spends great effort to assure information quality.

But what is information quality? – Up to now, there is no widely accepted common definition. Over the last 10 – 15 years miscellaneous studies were carried out – mainly by the MIT - to circumscribe the field of information quality. In late 2007 the *Deutsche Gesellschaft für Informations- und Datenqualität – dgiq* (German Society for Information and Data Quality) published a first definition of information quality for Germany [1]. This definition consists of 15 dimensions. With these dimensions, now a common understanding of information quality within and between organizations is possible.

But this definition has to be seen as a first step. It does not help very much, if everybody just understands, what information quality means. Furthermore organizations need help to institutionalize, maintain and further develop information quality. Here ISO/IEC 15504 can support. Combining 15504 part 2 [2] with the dimensions of information quality gives a powerful means to assess the capability of an organization to handle information quality.

However, this combination does not lead to a “classical” Process Reference Model (PRM) and Process Assessment Model (PAM). As the dimensions are closer to product components than to processes, the term “Process” should be substituted by “Dimension” which leads to a “DRM” and “DAM”. Applying DRM and DAM then indeed leads to processes, which ensure the organizational capability in handling information quality.

In section 2 of this article, a short overview on the development of information quality over the last 10 – 15 years is given. Section 3 then describes the 15 dimensions of information quality and the approach

behind it. Section 4 shows, how ISO/IEC 15504-2 can be applied to information quality dimensions, and delivers the DRM and DAM.

2. A short history of Information Quality

If one asks for the leading expert on information quality, the answer automatically points to Prof. Richard Y. Wang from the MIT. In his article [3] published 1996 Prof. Wang developed a hierarchical framework to structure dimensions of information quality. This framework was developed in two steps. In a first step, Prof. Wang asked data users, what they understand by quality attributes of data. This research leads to 179 attributes. In the second step the same user now had to prioritize the attributes on a scale from 1 (very important) to 9 (unimportant). Based on this prioritization the top 20 dimensions were determined as follows:

Rank	Dimension	Average
1.	Believability	2,71
2.	Value-added	2,83
3.	Relevancy	2,95
4.	Accuracy	3,05
5.	Interpretability	3,20
6.	Ease of understanding	3,22
7.	Accessibility	3,47
8.	Objectivity	3,58
9.	Timeliness	3,64
10.	Completeness	3,88
11.	Traceability	3,97
12.	Reputation	4,04
13.	Representational consistency	4,22
14.	Cost-effectiveness	4,25
15.	Ease of operation	4,28
16.	Variety of data and data sources	4,71
17.	Concise	4,75
18.	Access security	4,92
19.	Appropriate amount of data	5,02
20.	Flexibility	5,34

Over the next years, Prof. Wang performed several studies and evaluations to refine the dimensions. Additionally the dimensions were structured into 4 categories:

- Intrinsic Data Quality – to foster confidence in data correctness and objectivity
- Contextual Data Quality – to assure that data is relevant for the user at the point of use
- Representational Data Quality – to foster the understandability of data
- Accessibility Data Quality – to assure that the user knows how – and under which constraints – data can be accessed

This led to the following structure as published in 2002 [4]:

Intrinsic DQ	Contextual DQ	Representational DQ	Accessibility DQ
Believability	Value-added	Interpretability	Accessibility
Accuracy	Relevancy	Ease of understanding	Access security
Objectivity	Timeliness	Representational consistency	
Reputation	Completeness	Concise representation	
	Appropriate amount of data		

Table 1: categories by Wang

Since then this structure remained stable and was used as basis by the dgiq.

3. Information Quality Dimensions

Regarding the dimensions of Prof. Wang, some topics had to be further discussed. On the one hand not all dimensions were fully defined and explained – and therefore not free of interpretation. On the other hand some dimensions were overlapping.

In 2007 dgiq started a project to develop a set of dimensions that are free of overlaps and sufficiently explained to avoid interpretation.

Late in 2007 dgiq published a set of 15 dimensions which were based on the work of Prof. Wang and cleared out the problems described above. Another major change was made in the wording. “Data” was changed to “information” to make the dimensions independent from the use in IT.

3.1. Information Quality Dimensions

The 15 dimensions published by the dgiq are:

- *Accessibility* – information is accessible if it is retrievable on a direct way and by easy procedures
- *Appropriate Amount of Information* – information has an appropriate amount of information if the amount of available information satisfies the defined requirements
- *Believability* – information is believable if certificates show a high standard of quality and information mining and dissemination are performed with high effort
- *Completeness* – information is complete if it is not missing and is available for the selected process steps at the defined points in time
- *Concise Representation* – information has a concise representation if precisely the required information is displayed in an adequate and easy comprehensible format
- *Consistent Representation* – information has a consistent representation if information is consecutively displayed in the same modality
- *Ease of Manipulation* – information is easy to manipulate if it is easy to change and usable for different purposes
- *Free of Error* – information is free of error if it is concordant with reality
- *Interpretability* – information is interpretable if it is apprehended in the same, technical correct modality
- *Objectivity* – information is objective if it is strictly factual and non-judgmental
- *Relevancy* – information is relevant if it is necessary for the user
- *Reputation* – information has high reputation if source, transport media and processing system show high reliability and competence
- *Timeliness* – information is timely if it pictures the real attributes of an object promptly
- *Understandability* – information is understandable if it is directly understandable by the user and can be used for the users purpose
- *Value-added* – information is value-adding if it can lead to a quantifiable raise of a monetary end function

Even though this definition leaves room for further discussion, it gives a complete set of dimension which is now widely accepted in the German information quality community.

3.2. Information Quality Categories

Based on Prof. Wang’s work, dgiq defined 4 categories to structure the dimensions:

4. ISO/IEC 15504-2 and Information Quality

ISO/IEC 15504-2 “*identifies the measurement framework for process capability and the requirements for*

- (a) *performing an assessment;*
- (b) *Process Reference Models;*
- (c) *Process Assessment Models;*
- (d) *Verifying conformity of process assessments.”*

So the major focus of ISO/IEC 15504-2 is on processes. It is easy to see, that the dimensions above are no processes and may rather be described as “attributes” of information quality. On the other hand, if organizations want to assure information quality, they have to implement processes.

As there is no better approach than ISO/IEC 15504 an “ISO/IEC 15504-like” model for information quality was published in late 2007 [5] but this was more hands-on than delivering an approach which shows consistencies and inconsistencies to ISO/IEC 15504.

To close this gap, this section will discuss, which elements of ISO/IEC 15504-2 will be fulfilled by an approach based on information quality dimensions as defined above.

As mentioned before, a Process Reference Model as required by ISO/IEC 15504-2 is hard to develop based on the information quality dimension. This will be discussed and shown in the next subsection.

But, by using a slightly changed approach (Dimension Reference Model - DRM), it can be proved, that all other elements of ISO/IEC 15504-2 can be satisfied and therefore an Information Quality Capability Model (IQCM) will make sense – which will be shown in the other subsections.

4.1. From PRM to DRM

The requirements for a Process Reference Model (PRM) as defined in section 6.2. of ISO/IEC 15504-2 can be taken as fulfilled by the information quality dimensions as described below

- Domain
 - Information Quality
- Community of interest
 - All organizations that have to handle with or rely on high quality data and information
- Unique Process descriptions and identification
 - Given by the requirement for the dimensions that they are free of interpretation and overlaps

The problematic section of 15504-2 is 6.2.4., “Process descriptions”. This section defines that “the process descriptions in the Process Reference Model incorporate a statement of the purpose of the process which describes at a high level the overall objectives of performing the process, together with a set of outcomes which demonstrate successful achievement of the process purpose”.

The information quality dimensions are defined by one simple sentence. This sentence could either deliver the purpose or the outcomes, but not both at the same time. Every idea, to overestimate a dimension than it really delivers will lead to a pure theoretical approach without delivering any value.

One of these ideas may be to use the sentence for both, purpose and outcomes, which may lead to the following description for the e.g. dimension “Objectivity”:

Identifier	Objectivity
Purpose	Information should be strictly factual and non-judgmental
Outcome	(1) the organization proves that every information is strictly factual (2) the organization proves that every information is non-judgmental

Table 2: possible PRM approach

Even though this representation fulfills the requirements of ISO/IEC 15504-2 section 6.2.4 theoretically, it just doubles information (in purpose and outcome) to reach compliance to a standard. This won't give any additional value to the user and therefore should be avoided.

Another approach with the same problem may be to split information into identifier and purpose, which will lead to

Identifier	Objectivity
Purpose	Achieve objectivity of information
Outcome	(1) the organization proves that every information is strictly factual (2) the organization proves that every information is non-judgmental

Table 3: another possible PRM approach

All in all it has to be declared, that no useful approach can be developed that fulfills the requirements of an ISO/IEC 15504 process description and adds value for the user at the same time.

Another possible way may be to further describe and detail the dimensions. But as this has not been done by now (and will not be in near future) it only can be done by further interpreting the dimensions – which conflicts with the intent of the dimensions.

To find a satisfying solution, it is proposed to define a DRM – Dimension Reference Model. Based on the requirements of a PRM a possible definition for a dimension may be

“the dimension descriptions in the Dimension Reference Model incorporate a set of outcomes which demonstrate successful achievement of the dimension”

With this definition the duplicates between purpose and outcomes are resolved without losing content.

A proposal of dimension descriptions for all dimensions is given in the appendix.

4.2. From PAM to DAM

After defining the DAM, the definition of a PAM has to follow. Therefore the applicability of the process attributes has to be proven.

First of all, for all process attributes and further definitions “process” has to be changed to “dimension”.

For process attributes (PA) 3.1 to 5.2 it is easy to see that these are applicable to the DRM. As PAs 3.1 and 3.2 cover standardization and deployment of the related objects, PAs 4.1 and 4.2 measurement and control and PAs 5.1 and 5.2 innovation and optimization all are applicable to processes as well as to dimensions.

PA 2.2 is also applicable to dimensions as objectives can be defined for each dimension, the performance of handling the dimension can be planned, monitored and adjusted, responsibilities for each dimension and the related outcomes can be defined and resources and interfaces can be managed.

In the very end, only PA 1.1 and PA 2.2 require further treatment. PA 1.1 addresses process purpose. As there is no dimension purpose defined in the dimension description, PA 1.1 has to be redefined as follows:

PA 1.1 Dimension performance attribute

The dimension performance attribute is a measure to which extend the requirements of the dimension achieved. As a result of full achievement of this attribute

a) the dimension achieves its defined outcomes

The second PA to be investigated is PA 2.2 – work product management attribute. For handling the dimensions the usual work products are evidence documents that prove the achievement of the dimension outcomes.

But, even if the first sentence of the PA definition is changed, the rest of the PA (concerning requirements for work products, documentation and control, work product identification, documentation, controlling, reviews and adjustment) will still be valid.

A possible basic definition of PA 2.2 for dimensions therefore may be:

PA 2.2 Work product management attribute

The work product management attribute is a measure of the extent to which the evidence documents proving the achievement of the outcomes of the dimension are appropriately managed.

....

With these changes, a consistent DRM with an applicable DAM exists.

4.3 ...and the rest?

ISO/IEC 15504-2 is more than defining requirements for a PRM and PAM. Therefore the other elements should now be further analyzed to show the applicability of ISO/IEC 15504-2 in information quality.

Section 4 of ISO/IEC 15504-2, concerning the performance of an assessment, is applicable as long as the PRM is changed to DRM and therefore the PAM has to become a DAM. All other elements like the assessment process, roles and responsibilities, defining the initial assessment input and recording the assessment output can be used for DRM and DAM.

The measurement framework, as described in section 5 of ISO/IEC 15504-2 was already widely discussed in section 4.2 of this article. The only thing that is missing is the attribute rating (section 5.7 of ISO/IEC 15504-2), but the rating scale can be used within the DAM and therefore section 5 of ISO/IEC

15504-2 is fulfilled if the PAs 1.1 and 2.2 adapted as described above.

The requirements of section 6 (Models for process assessment) are fulfilled by DRM and DAM as long as the constraints as described above are taken into account

Section 7 of ISO/IEC 15504-2, defining the mechanisms for verification of conformity to ISO/IEC 15504 is seen as out of scope for the Information Quality Capability Model, as the major goal of this model is not conformance to ISO/IEC 15504, but giving a means for institutionalizing processes to reach high information quality.

5. Conclusion

The DRM and DAM as defined in this document are not fully compliant – but this is not the major goal! As the model based on DRM and DAM delivers a methodology to assure and assess the quality and capability of an organization to handle and improve information quality, the major goal from the point of view of the information quality community is reached.

However, a full alignment between information quality dimensions and ISO/IEC 15504-2 can be reached in the future. Therefore the dimensions have to be further detailed. If the current dimensions are seen as purpose, a detailing of the dimensions would automatically lead to outcomes which have to be facilitated by base practices.

So the next step is on the information community to deliver more detailed information on information quality dimensions to reap the benefits of the well established ISO/IEC 15504 standard.

6. References

- [1] dgiq: IQ-Definition. Retrieved January 9th, 2007 from www.dgiq.de/_rubric/index.php?rubric=IQ-Definition
- [2] ISO/IEC 15504-2:2003, *Information Technology – Process Assessment – Part 2 Performing an assessment*, ISO/IEC, 2003.

[3] Wang, R. Y., Strong, D. M.: "Beyond Accuracy: What Data Quality Means to Data Consumers". *Journal of Management Information Systems* Vol. 12, No. 4, pp. 5-34, 1996.

[4] Kahn, B. K., Strong, D. M., Wang, R. Y.: "Information Quality Benchmarks: Product and Service Performance". *Communications of the ACM*, Vol. 45, pp. 184-192, 2002.

[5] Malzahn, D.: *Informationsqualität bewerten*, Fachbuchverlag Richartz und Kurpicz, Lünen, 2007.

7. Appendix

Identifier	Accessibility
Outcome	(1) The organization proves that information is retrievable on a direct way (2) The organization proves that information is retrievable by easy procedures

Identifier	Appropriate Amount of Information
Outcome	(1) The organization proves that the amount of available data satisfies the defined requirements

Identifier	Believability
Outcome	(1) The organization proves that certificates show a high standard of quality (2) The organization proves that information mining is performed with high effort (3) The organization proves that information dissemination is performed with high effort

Identifier	Completeness
Outcome	(1) The organization proves that no information is missing (2) The organization proves that information is available for the

	selected process steps (3) The organization proves that information is available at the defined points in time
--	---

Identifier	Concise Representation
Outcome	(1) The organization proves that precisely the required information is displayed in an adequate format (2) The organization proves that precisely the required information is displayed in an easy comprehensible format

Identifier	Consistent Representation
Outcome	(1) The organization proves that information is consecutively displayed in the same modality

Identifier	Ease of Manipulation
Outcome	(1) The organization proves that information is easy to change (2) The organization proves that information is usable for different purposes

Identifier	Free of Error
Outcome	(1) The organization proves that all information is concordant with reality

Identifier	Interpretability
Outcome	(1) The organization proves that information is apprehended in the same, technical correct modality

Identifier	Objectivity
Outcome	(1) the organization proves that every information is strictly factual (2) the organization proves that every information is non-judgmental

Identifier	Relevancy
Outcome	(1) The organization proves that the information is necessary for the user

Identifier	Reputation
Outcome	<p>(1) The organization proves that the source of information shows high reliability and competence</p> <p>(2) The organization proves that the transport media for information shows high reliability and competence</p> <p>(3) The organization proves that the system for processing information shows high reliability and competence</p>

Identifier	Timeliness
Outcome	(1) The organization proves that information pictures the real attributes of an object promptly

Identifier	Understandability
Outcome	<p>(1) The organization proves that information is directly understandable by the user</p> <p>(2) The organization proves that information can be used for the users purpose</p>

Identifier	Value-added
Outcome	(1) The organization proves that information leads to a quantifiable raise of a monetary end function