A Management System for Data Quality

Towards a Solid, Coherent and Standardised Approach

DAMA-NL Data Quality Working Group:

Peter van Nederpelt, Kamal Ahmed, Andrew Black, Marco Heij, Aris Prins, René Wiertz, Dirk Hueting, Laurens van der Drift

data quality, factsheet, ISO 9001, Data Quality Management System

Introduction

In the data domain, ISO 27001 *(Information security management systems - Requirements)* is a well-known standard in the field of information security. This standard is widely used, and organisations may be certified as conforming to it. This is a means of showing the outside world that they have their information security in order.

There are comparable standards for data quality, namely ISO 8000 *(Data quality)* and ISO 9001 *(Quality management systems – Requirements),* but the former is applied in too specific a context and the latter is rarely applied to data quality. An important reason for this in the case of ISO 9001 is that this standard is generic and not specifically designed for application to data. ISO 9001 is intended for the quality of products and services in general.

This white paper describes an approach to applying ISO 9001 to data quality. The terms for the elements used in ISO 9001 are converted to terms that are common in the data domain. The meaning of each of these elements is then briefly explained. The aim of this is to make it easier to apply ISO 9001 to the management of data quality.

Management system

In the data domain, people generally talk about data management and data governance. ISO uses a different concept, namely that of a management system. This is used in both ISO 27001 and ISO 9001. A management system is seen as a set of interrelated or mutually influencing elements aimed at determining the objectives of an organisation and achieving these objectives.

The ISO standards are updated every five years following a worldwide consultation. Certification for an ISO standard can be done by an external auditor accredited for this purpose. A management system can relate to various disciplines, such as quality management and information security, but also to environmental management and financial management. The ISO standards for management systems contain a common core: the high-level structure. This is because the various management systems have things in common.

Business drivers for implementing a Data Quality Management System (DQMS)

Reasons for introducing a DQMS are that an organisation wants to achieve and maintain good data quality. This may be necessary, for example, to enable the use of data for multiple purposes. In the first place, data can be used in the regular processes, such as purchasing, production, sales, finance, and HR. Insufficient data quality in these processes leads to inefficiencies and irritations, both internally and externally.

In the second place, data can be used for reporting to regulators and society, for data-driven work, for management based on key performance indicators (KPIs) and for training, evaluating, and improving machine learning systems within the domain of Artificial Intelligence (AI). The use of data of insufficient quality can lead to incorrect decisions, resulting in fines, life-threatening situations for individuals or groups, and loss of reputation.

In the financial domain, it is normal to actively manage the quality of financial data. After all, financial statements must be in order and shareholders must be able to rely on them. Furthermore, organisations' annual financial statements are audited to ensure the quality. Within this domain, it is customary to have the financial management system in order in the form of the administrative organisation (AO).

Alternative frameworks

In the world of data management there are frameworks that are similar to the DQMS explained in this paper. First, the standard ISO 8000 has existed for quite some time. This consists of 13 parts but has a mainly technical character and focuses on data quality of master and reference data. Then there is the DECAM system. This system is aimed at data management as a whole and thus has a broader scope. DAMA DMBOK is a body of knowledge and cannot really be called an assessment framework. A DQMS is completely focused on data quality.

Factsheets

The Data Quality working group of DAMA Netherlands has set itself the goal of describing all elements of the DQMS in clear documents, also known as factsheets. The factsheets are published on the DAMA-NL website (www.dama-nl.org/datakwaliteit) in the data management wiki. The wiki format has been chosen because of the accessibility and maintainability of the factsheets.

A factsheet consists of several sections: a definition of the element, the purpose of the element, relationships with other elements, the PDCA cycle (Deming) of an element, and the properties of an element. Also, examples are given, and a story is told, showing how the element can be applied. The description of the element is taken from various sources such as DAMA DMBOK2, ISO 9001, and books and articles that can be found on the Internet. See the list of references below.

Specifications of requirement of each element

In this paper we will not go so far as to specify requirements of each element of the management system, i.e., what are the requirements of a (data) quality policy. This is also not necessary here, as these requirements are already contained in the ISO 9001 standard.

Certification

If sufficient attention is paid to all elements of the Data Quality Management System, it is possible for an organisation to be certified for ISO 9001 by an external, certifying body (external audit), even if it concerns data quality. You then limit the scope of the certificate to the data processing processes.

Selection and implementation sequence

This paper does not provide advice on which elements are relevant for an organisation or in what order they should be implemented. This is highly dependent on the environment in which they are applied, the resources available and the maturity level of the organisation. Incidentally, it is required for certification for ISO 9001 that all elements are in force but an implementation order is not indicated. This is up to the organisation itself.

PDCA cycle of the elements

We assume that there is a PDCA cycle for each element. That is, an element is prepared ("Plan"), implemented ("Do"), evaluated ("Check") and, if necessary, improved ("Act"). All elements are therefore in an improvement cycle. The cycle will therefore not be repeated for each element in this paper.

Objectives and elements of the DQ Management System (DQMS)

Table 1 lists all the objectives and elements of the DQMS and translates the terms from ISO 9001 into terms that are common in the data domain.

Table 1 Illustration of elements of ISO 9001 on the DQMS

ISO 9001 paragraph	ISO 9001 Elements	2 Objectives and 28 Elements of the DQMS
4.1	Understanding the organisation and its context	Context analysis
4.2	Understanding the needs and expectations of interested parties	Stakeholder analysis Compliance
4.3	Determining the scope of the quality management system	Scope
5.1	Leadership and commitment	Leadership
5.2	Policy	Data quality policy
5.3	Organizational roles and responsibilities	Roles and responsibilities
6.1	Actions to address risks and opportunities	Risk analysis
6.2	Quality objectives and planning to achieve them	Data quality objectives
7.1	Resources	Resources
7.2	Competence	Staff competence
7.3	Awareness	Awareness of data quality
7.4	Communication	Communication about Data Quality
7.5	Documented information	Metadata
7.5.3	Control of documented information	Document control
8.1	Operational planning and control	Data processes
8.2	Requirements for products and service	*Objective:* Meeting data quality requirements
		Critical data elements
		Data quality rules
8.3	Design and development of products and services	Data design processes
8.4	Control of externally provided processes, products, and services	Data suppliers
8.5.2	Identification and traceability	Data lineage
8.7	Control of non-conforming outputs	Data cleansing Data issues (1)
9.1	Monitoring, measurement, analysis, and evaluation	Data profiling Monitoring Data Quality
9.1.2	Customer satisfaction	*Objective:* Data consumer satisfaction
9.2	Internal Audit	Internal Audit
9.3	Management review	Management review
10.2	Non-conformity and corrective action	Data issues (2)
10.3	Continual improvement	Continuous improvement

Objectives of the DQMS

A DQMS has the two objectives below and they are also an extension of each other: objective 1 also contributes to objective 2.

1 Data quality requirements

The first objective is to meet the requirements of data quality. Data quality has many dimensions. The DAMA NL Data Quality working group has identified 60 of them and provided them with a standardised definition. Of course, a choice has to be made from these. This also applies to the data to which these dimensions relate. The 'critical data element' can play a role here.

2 Data consumer satisfaction

The second objective is satisfaction of the data consumers: those who use the data inside and outside the organisation. The perception of the data consumers about data quality is measured in one way or another. Conclusions are drawn from these measurements and, if necessary, actions are planned and implemented.

Elements of the DQMS

In this section, all elements of the DQMS are briefly explained. The elements are described in the sequence of the metro model in figure 1. For a more detailed, structured description of the elements, please refer to the DAMA-NL factsheets in the data management wiki (https://datamanagement.wiki/overview/overview_data_quality_management_system).

1 Compliance
Compliance with laws, regulations, or standards is the degree to which data is in accordance with laws, regulations, or standards.

2 Data quality objective
With quality objectives, the requirements that the data must meet are formulated in a measurable way. This is then also measured and tested against a standard. If the standard is not met, improvement actions are required. This is the dashboard about data quality.

3 Data quality policy
The data quality policy states what intentions the management has with regard to data quality. For example, which dimensions of data quality are important and which data are involved. It can also state which legislation and regulations apply to data quality. Data quality policy gives direction to actions in the field of data quality.

4 Internal audit
An internal audit examines on behalf of the management, whether all elements of the DQMS are demonstrably present and whether they are effective. The results of an internal audit are reported to the management. Based on the results, the management determines whether improvements to the system are required.

5 Management review
Management periodically assesses whether the DQMS is sufficiently effective and whether it needs improvement. It does this based on various reports and documentation. This makes management responsible for the system. This responsibility cannot be delegated.

6 Leadership

A DQMS does not get off the ground without management involvement. It requires a bottom-up approach as well as a top-down approach. Management is also directly involved in the management system through the management review.

7 Scope
The area over which an activity extends.

8 Stakeholder
A stakeholder analysis examines who is involved in the data processing process and what interests, expectations and perceptions they have about data quality. Stakeholders can be found on the input side (suppliers), the throughput side and the output side (internal and external data consumers). In managing data quality, account is taken of the influence that stakeholders have on data quality and on the impact that data quality has on their functioning.

9 Data supplier
Data suppliers can consist of organisations that deliver files, but also users from inside and outside the organisation who register data can be considered as data suppliers. Suppliers are assessed on their provision of input of sufficient quality. If this is not the case, measures must be taken.

10 Data quality communication
Communication on data quality aims to make stakeholders aware of the importance of data quality and their role in it. Through communication, it is also made clear what improvements have already been made and what the planned improvement actions are.

11 Roles and responsibilities
Roles and responsibilities mean that it is clear who has which roles and responsibilities with regard to data quality. Roles that are often mentioned are the Chief Data Officers, the data owner and the data steward, but other employees also play a role in data quality.

12 Competence of staff
Employees involved in data processing should be supported by management and HR in acquiring the knowledge and skills to be able to perform their tasks properly. It is clear that if employees are insufficiently competent, this will have consequences for data quality.

13 Awareness of data quality
Awareness of Data Quality is a recognition or awareness by everyone in the organisation that only with the right skills and competences, data of sufficient quality can be realised. They are aware that there is a quality policy, that there are quality objectives, and what their contribution is to an effective DQMS.

14 Documented Information
Information required to be controlled and maintained by an organization including the medium on which it is contained.

15 Risk analysis
A risk analysis investigates which events or causes could threaten the objectives of the system (sufficient data quality, satisfaction of data consumers and data producers). It also describes what measures have already been taken to reduce this risk. Finally, it is assessed whether the measures are adequate or whether additional measures are required.

16 Context analysis
In a context analysis, the most important internal and external issues about data quality are examined. This is important to prioritize improvement actions, to focus on the main problems and to anticipate

future developments.

17 Improvement Action
Continuous improvement should not only ensure that all elements of the management system are periodically evaluated and improved. There must also be room for spontaneous ideas for improvement. That is, there must be a working method that supports this.

18 Data Model
Data model ensure that the right data is recorded. For this purpose, the information needs of data consumers are identified. In addition, the data is put into the right structure.

19 Data process
Data processing processes ensure that input is processed into output with the correct data quality. Output can also consist of a database that is consulted by employees or systems.

20 Resource
Resources to consider when developing a data quality business case to cleanse data and improve quality include people, applications, data platforms, and budget.

21 Metadata
Metadata describes the meaning of data. Without metadata, the quality of data cannot be determined. Also, without metadata, data is meaningless to data consumers. Metadata is correct, complete and clear. It is also maintained.

22 Data quality rule
Data quality rule is a rule with which data must comply. If they are implemented in the systems, they can prevent errors. Data quality rules must also be managed. They can be incomplete or incorrect. This can lead to a data quality leak.

23 Critical Data Elements
Critical Data Elements are attributes that are essential to the organisation. This can also be expressed in gradations. When managing data quality, these are the attributes or fields that should be considered first. If it is more convenient to consider certain data sets as critical, then of course there is no objection to that (critical data sets).

24 Data lineage
Data lineage is metadata that describes the origin of data. Via this metadata, it is possible to trace the source of certain output and how the data has been processed. Sometimes this is important, to be able to justify yourself as an organisation to an authority. This metadata can also be important when adjusting or redesigning data processing processes. These may not lead to a decline in data quality.

25 Data cleansing
In data cleansing, data issues are actively traced and subsequently resolved. The DQMS has a working method for carrying out data cleansing, whether or not with the help of a tool.

26 Monitoring data quality
A monitoring process ensures that the quality of the data is continuously monitored. This will mainly take place on the basis of data quality rules. If the development is made visible in a timeline, trends can be identified, and timely action can be taken.

27 Data quality issue
The handling of data issues is an important and concrete element within a DQMS. If an employee encounters one or more data issues, he can report them. These are then assessed and dealt with

('correction'). The way they are dealt with is also registered and reported back to the reporter. Periodically, an assessment is made to determine whether a pattern can be recognised in the data issues and whether measures can be taken to prevent the issues ('corrective actions').

28 Data quality analysis

Data quality analysis precedes data cleansing. In data quality analysis, the quality of the data is illustrated by means of statistics. Statements are made in figures about entire data collections. Based on this, it can be decided which data is eligible for quality improvement. Data quality analysis can also be used to monitor data quality.

Metromodel

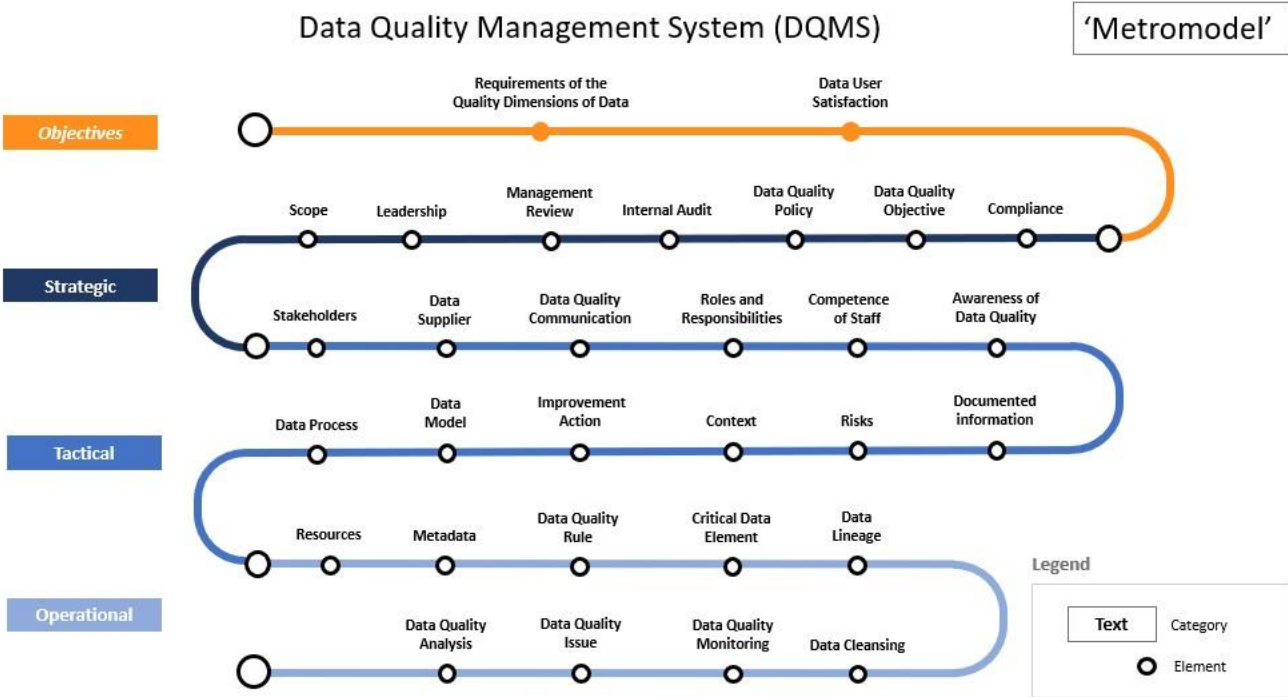Figure 1 summarises the objectives and elements of the DQMS once again in the form of the Metromodel.

Figure 1 Objectives and elements of the DQMS

Conclusion

It is important that an organisation has its data quality in order. Applying a management system in this area can certainly help. ISO 9001 offers all the tools necessary to do this in a standardised manner. An organisation can also have itself certified by a certifying body (external auditor), so that it can show the outside world that the data quality is being properly managed. We can conclude that it is quite possible to apply the concept of a management system to data quality, despite the fact that different jargon is used.

Finally

Further comments on this document are highly appreciated by the DAMA-NL Data Quality Working Group. Reactions can be sent to WGDQ@planetbox.nl. The workgroup is particularly interested in which questions the DQMS raises, which considerations can hinder the introduction of such a framework, and which potential the DQMS has for various organisations.

References

DAMA (2017). DAMA-DMBOK. *Data Management Body of Knowledge.* 2nd Edition. Technics PublicationsLlc. August 2017.

DAMA Dictionary of Data Management. 2nd Edition 2011. Technics Publications, LLC, New Jersey.

Data quality. (2021, December 4). DAMA Foundation. https://www.dama-nl.org/datakwaliteit

DECAM. About DCAM – EDM Council

Factsheets from DAMA-NL. https://datamanagement.wiki/

ISO 27001:2017. *Information technology – Security techniques – Information security management systems – Requirements*

ISO 8000. *Data Quality* 13 parts.

ISO 9000:2015. *Quality Management Systems - Requirements.*

ISO 9001:2015. *Quality Management Systems - Fundamentals and vocabulary.*

Version 2.4 EN. Dated 11 December 2024