Awareness of data quality

Definition

Awareness of data quality is an acknowledgement or realisation by everyone in an organisation, that their current skills and competencies are (becoming) ineffective, resulting in data of inadequate quality.

Note 1

Awareness is attained when people understand their responsibilities and how their actions contribute to the achievement of the organisation’s objectives (ISO 9000:2015).

Knowledge is generally, expertise; familiarity gained through experience or association; recognition of a situation and familiarity with its complexity. Understanding of the significance of information. (DAMA-DMBOK Guide, 1st edition, page 3.)

Situational awareness is the perception of an environment's state and conditions at a point in time. (DAMA-DMBOK Guide, 1st edition).

Note 2

ISO 9001 states, that the organisation shall ensure that persons doing work under the organisation's control are aware of (ISO 9001:2015):

1. The data quality policy.
2. Relevant data quality objectives.
3. Their contribution to the effectiveness of the data quality management system, including the benefits of improved performance.
4. the implications of not conforming with the data quality management system requirements.

Note 3

In practice, skills and competencies are very often used interchangeably, but are different:

	Definition	**Examples**
Skill	Specific learned abilities one requires to perform a given task successfully	Programming-skills, problem-solving skills
Competency	Knowledge and behaviours that lead one to be successful at a task	Conscientious, accurate, knowing about standards and audits applicable in certain sectors

Note 4

Awareness and knowledge are also used interchangeably, whereas in training & development, these

are considered as categories or labels for proficiency, each with their own specific levels for proficiency:

Proficiency Category	**Proficiency Level**	**Skill**	**Competency**
Mastery	Level 4.x Level 4.1	(out of scope for this factsheet)	(out of scope for this factsheet)
Skilled	Level 3.x Level 3.1	(out of scope for this factsheet)	(out of scope for this factsheet)
Knowledge	Level 2.x Level 2.1	Skills to e.g. inspect and correct data quality issues using tools with some help or guidance from others.	Conscientious, accurate approach, etc
Awareness	Level 1	Skills to acknowledge, recognise or become familiar with certain terminology, procedures, policies, other people's work, etc, as a foundation to improve data quality.	General interest in data quality developments, challenges, or having a "helicopter view", etc.

Awareness of data quality has only one proficiency level and is about leadership and staff having the skills and competencies to recognise or acknowledge the occurrence of inadequate data quality and its impact on the organisation.

After having awareness, the next step is "Knowledge of data quality". This is about having the skills and competencies to take effective actions to improve the quality of data to an adequate

level with some guidance from others.

The labels of the proficiency categories "Skilled", "Mastery", the number of proficiency levels, and descriptions of the skills and competencies may vary from organisation to organisation.

Synonyms

- Data quality cognizance
- Data quality consciousness
- Data quality familiarity
- Data quality knowledge
- Data quality realisation
- Data quality understanding

Purpose

The purpose of awareness of data quality from leadership and staff, is to positively contribute to:

1. Data of adequate quality
2. Customer experience,
3. Revenue and costs,
4. Morale,
5. Organisational effectiveness,
6. Organisational competitiveness,

7. Decrease risk levels of inadequate data quality.

Life cycle

Phase	Activity
Plan	* To measure awareness * To prepare awareness program
Do	* To execute awareness program (raise awareness) * To repeat awareness program periodically
Check	* To monitor awareness* To evaluate awareness program
Act	* To adapt awareness program

Note 1

Managing awareness of data quality may have different objectives (i.e. creating awareness, maintaining awareness, or increasing awareness):

1. Activities related to creating awareness for Data Quality if there is no awareness at all from the start. This situation would typically occur in a 'green-field' situation where no prior initiatives have taken place to assess the awareness of Data Quality among staff and leadership in an organisation.
2. Activities related to maintaining awareness levels for Data Quality, if the awareness levels have (suddenly) dropped and do not align with an agreed ambition level for data quality.
3. Activities related to increasing the awareness levels for data quality to reach a higher ambition level for Data Quality.

When mentioning 'improving awareness' or 'raising awareness', the above objective needs to be clearly stated, as well as for whom (leadership and/or staff), over which period of time and through what medium (i.e. trainings, interactive workshops, awareness campaigns, internal audits, coaching, etc).

Note 2

Improving or raising awareness through communications requires the communications to:

- be of the right level,
- be of the right format, and
- occur at specific moments and not too frequently

This will allow the intended audience to internalise the awareness message and reduce the risk of ignoring the message.

Managing Awareness of data quality through monitoring & evaluation

- Plan for recurring in-person trainings, or instructor-led online trainings every 6 months to raise awareness of data quality. In these trainings, use gamification or serious gaming to initiate discussions between employees and leadership on data quality pain points for the business. These trainings could be facilitated by Learning & Development, or by Data Stewards.
- Capture learnings from these trainings in the form of documentation on "practices worth replicating", "best practices", or "bad practices", and embed these in Knowledge Management

systems. Update these documents as frequently as needed over time.
- Develop just-in-time learning solutions by enabling learners to access information on data quality using a Knowledge Management system. Learners can search for information on data quality when they need this, from any device, and from any location.
- Evaluate the effectiveness of learning interventions for the business through Kirkpatrick Foundational Principles and using the Kirkpatrick levels of training evaluation. Evaluations can be used to improve/adapt trainings to the required competencies.
- Evaluate the effectiveness of learning interventions for the business by monitoring the data quality before and after every learning intervention, for example by monitoring data quality indicators in a dashboard.
- Show the data quality scores in relation to meeting regulatory compliance, certification, safety levels, number of casualties reported, up-time of assets, etc. using a dashboard that is accessible to everyone.
- Reactive improvement of data quality versus pro-active improvement depends on a costbenefit analysis to mitigate risks. What are the risks for reputational damage, business continuity, safety, environmental impact, etc, if scheduling data quality improvement activities are postponed, performed at fixed time intervals only, or on the basis of predictions?
- Consider data as an ‘asset’, similarly to ‘asset management’, may follow the ‘As Low As Reasonably Practicable’ (ALARP) risk mitigation framework.
- Consider developing, documenting, monitoring and enforcing Service Level Agreements (SLAs) for data as an asset. SLAs would be for example: availability of the data, timeliness of the data, integrity of the data as part of performing a data migration, etc.
- Awareness of Data Quality is a ‘must’ for ethical Artificial Intelligence. Data for, and from machine learning models, should be evaluated along dimensions such as Fairness, Accountability, Confidentiality and Transparency (FACT).

Characteristics

Characteristic	Requirement
Effectiveness of awareness of data quality	Leadership and staff understand the consequences of inadequate data quality of the organisation. Their understanding meets the required proficiency level for awareness. Leadership and staff attained the required proficiency level, which results in data of adequate quality.
Cost-effectiveness of awareness of data quality	Awareness of data quality leads to a positive business case, i.e., the benefits are higher than the costs.

Relations

- Awareness of data quality is guided by leadership.
- Awareness of data quality is improved by communication about data quality.
- Awareness of data quality is improved by internal audit.
- Awareness of data quality is guided by data quality objectives.
- Awareness of data quality evaluates **Staff competence**.
- Awareness of data quality guides data quality policy.
- Awareness of data quality acknowledges roles & responsibilities.
- Awareness of data quality applies risk analysis.

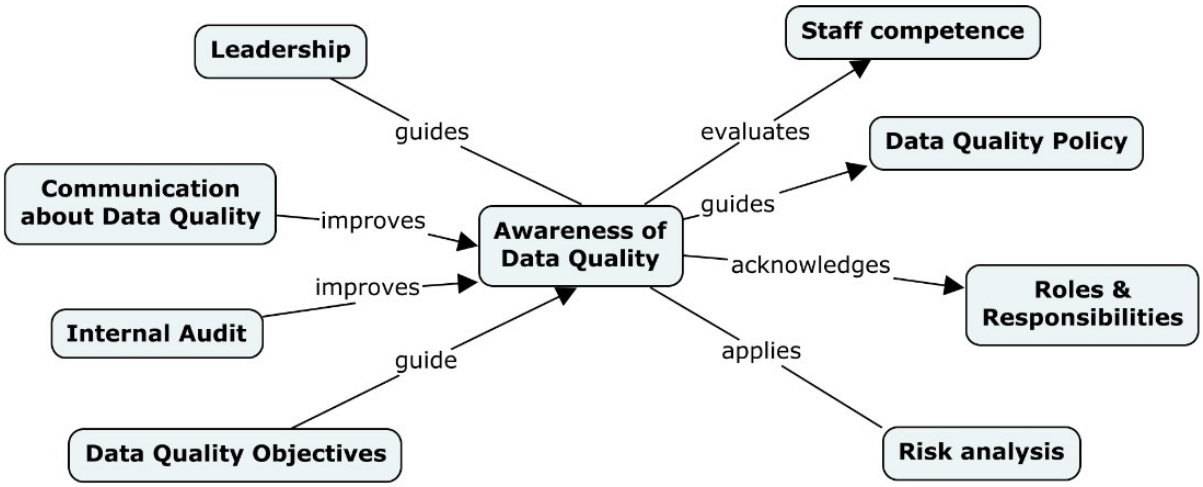

Figure 1: Relationships of Awareness of Data Quality with other concepts.

Stories

Below are three stories set in three different contexts, high-lighting the occurrence of awareness and its implications for meeting an overall objective. Text shown in **bold** indicates what the awareness means, and the purpose is shown in *italic*.

Story 1: Being aware of signs related to a cardiac arrest and providing adequate first-aid

One day you decide to take a brisk walk in a park. Just before reaching home, you see a pedestrian collapsing at the corner of the street. You quickly rush towards the pedestrian to see if you can *help*. Having been trained in basic first-aid, **you are able to quickly recognise the key-signs of a cardiac arrest**. You immediately call for an ambulance and explain the situation to the operator. **Acknowledging, that you have not been trained to provide mouth-to-mouth resuscitation, you suddenly remember that an Automated External Defibrillator (AED) may be of help**: AEDs can assist users to help stabilise patients. You ask a bystander to locate an AED device. While attaching the AED device to the patient's body, the ambulance responders arrive. Several days after this incident, **you decide to take an advanced course in first-aid**.

Story 2: Using awareness campaigns to reduce 'drink & drive traffic casualties'

Consuming alcohol now and then may be considered a relaxing pastime. Driving a car can be comfortable, quick and safe. But: driving shortly after consuming alcohol may cause unsafe situations for all traffic participants, possibly causing traffic casualties. Governments can use communication campaigns specifically aimed at youngsters. These campaigns can **help youngsters to change their behaviour towards drinking and driving**, so that *the number of traffic casualties will reduce*. The awareness message may, for example, focus on how alcohol affects youngsters' ability to safely control a vehicle.

Story 3: Being aware of how the quality of data impacts the overall quality of a report

An organisation in the financial sector had been collecting data on separate storage repositories for several years, without regularly looking at the quality of the data. To simplify the annual reporting process, managers of various business-units suggested to first migrate all data to a central database and to improve any issues with data quality after the migration. One manager suggested the opposite: performing a data migration has risks and requires a careful planning and execution, so why not first understand *how the quality of the data impacts the final report*? Over the next few weeks, she prepared and organised several interactive workshops with leaders and representative staff of the business-units. **She could relate to how participants grappled with interpreting the data quality levels in their business-units**. She then compared these results with results for data-lineage and presented this to the entire team. **Now everyone became aware** of *how the quality of the data was impacting the overall quality of the reported data*.

References

ALARP (As Low as Reasonably Practicable), Wikipedia article.

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

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

DAMA-DMBOK Guide, 1st edition.

DAMA Netherlands Working Group Data Quality.

Fairness, Accountability, Confidentiality, and Transparency (FACT) in AI, A. Lucic, 9 April 2020, University of Amsterdam.

ISO 9000:2015(en). Quality Management Systems – Requirements.

ISO 9001:2015(en). Quality Management Systems – Fundamentals and vocabulary.

Overview of the Kirkpatrick Model and foundational principles, April 2020.