Defining metadata

Definition

Definining metadata are metadata that enables interpretation of the meaning of data to provide information.

Purpose

The purpose of defining metadata is to enable consistent interpretation of data and thereby provide information. Which comes first, the data or the metadata? In principle, the defining metadata is created in the data design phase as specification of the implementation and is maintained as reference source for users of the data.

Life cycle

Phase	Activity
Plan	• Compose defining metadata • Establish defining metadata
Do	• Use defining metadata
Check	• Evaluate defining metadata
Act	• Revise defining metadata

Characteristics

Characteristic	Requirement
Accuracy of defining metadata	Defining metadata should be accurate enough.
Completeness of defining metadata	Defining metadata should be complete.
Unambiguity of defining metadata	Defining metadata must not be open to misinterpretation
Clarity of defining metadata	Defining metadata should be legible and understandable

Relations

- Definining metadata is subtype of metadata.
- Inputs to and outputs from defining metadata are shown in Figure 2.
    - Data quality requirements, data quality policy and design processes are input to defining metadata.
    - Data quality policy and Requirements provide principles and policies for the creation and management of defining metadata. * Design processes provide methodologies for the creation and management of defining metadata.
- Defining metadata is input to
    - Data quality objectives,
    - Data quality policy,
    - Data quality rules,
    - Data quality monitoring,
    - Data issues,

    - Awareness of data quality, and
    - Data cleansing.
- Data quality objectives can include requirements for availability and application of Defining Metadata.
- Defining metadata are a basis for determining data quality rules.
- Defining metadata provide standards for data quality monitoring.
- Defining metadata provide criteria when resolving data issues.
- Defining metadata assist in making awareness of data quality explicit.
- Defining metadata provide a basis for the selection of critical data elements, and
- Defining metadata provide criteria to be applied in data cleansing.

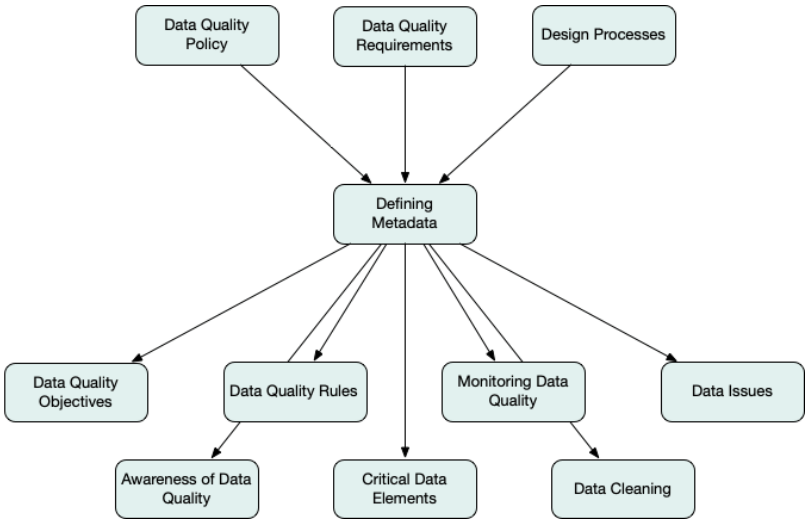

Figure 2 Inputs to and Outputs from Defining Metadata

An architecture of defining metadata

Figure 3 shows the three levels of defining metadata, conceptual, logical, and technical with the main forms of metadata at each level and their relationships. The diagram also indicates the roles that are primarily engaged with the metadata at each level.

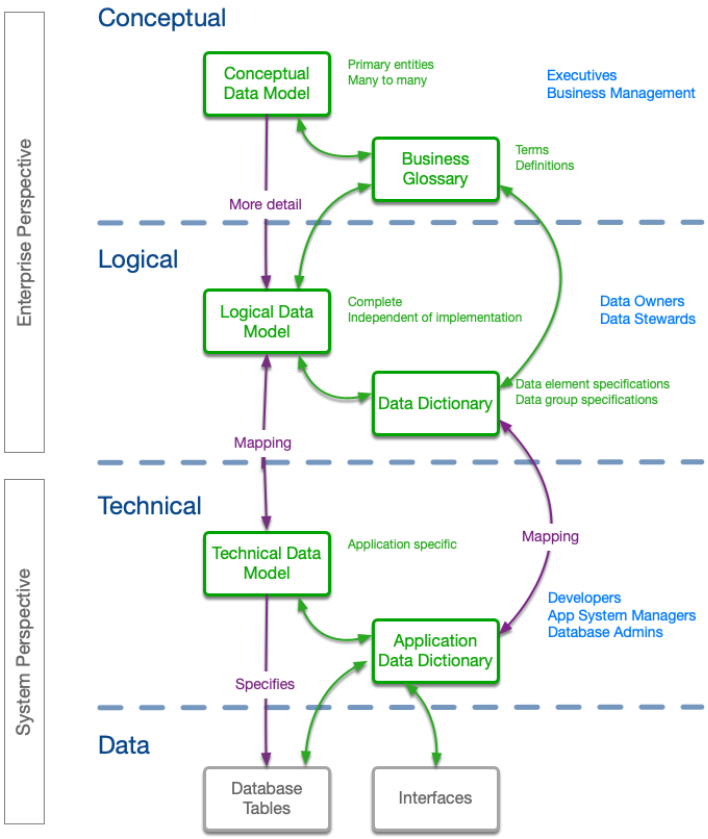

Figure 3 An architecture of defining metadata

Some forms of defining metadata

Data models

The goal of semantic modelling is the creation of a common understanding of the meaning of things, thereby helping people understand each other and done in such a way that the meaning is explicit and accurate and is understood by humans and interpretable by computer systems.

Although the terminology varies from one methodology to another, the elements to be found in most semantic modelling languages are: entities, relations, classes, attributes, terms, and axioms.

Examples of data element specifications

Identifier:	DE001
Name:	Code of a customer as DUNS Number
Definition	Code identifying a customer according to the Data Universal Numbering System (DUNS) of Dun & Bradstreet.
Data type:	Numeric
Format:	Fixed length 9
Value domain:	DUNS code numbers issued by Dun & Bradstreet https://www.dnb.com https://www.altares.nl

Identifier:	DE002
Name:	Loaded weight of a shipping container
Definition	The weight of a shipping container including its contents according to a weight unit of measure.

Data type:	Real number
Format:	Variable length maximum eight digits with two decimal positions
Note:	This data element type requires an associated code of a weight unit of measure.

Identifier:	DE003
Name:	Code of a unit of measure UN/ECE Rec. 20
Definition	Code identifying a unit of measure according to UN/ECE Recommendation 20
Data type:	Alphanumeric
Format:	Variable length maximum three characters
Value domain:	Code of a unit of measure according to Recommendation No. 20 Codes For Units Of Measure Used In International Trade published by United Nations Economic Commission For Europe (UN/ECE) https://unece.org/trade/uncefact/cl-recommendations

Such data element types are registered and maintained in a data dictionary. They are related to data models as specifications of the implementation of attributes of entities.

The names and definitions of data element types are used in user interfaces and example values may be registered in the data dictionary for use as prompts in electronic forms that are filled manually.

Example DE001 would be part of customer master data. It might be the primary identifier of a customer record or may be a secondary identifier used in credit check processes or as a means of building the hierarchy of related corporate organisations.

Example DE002 would be used during the trajectory of a shipment, for example when placing an order for shipment of a container, in the loading plan of containers in a vessel and in declarations to authorities.

Story

For some time issues and arguments had been rumbling on about who was responsible for recurring errors and delays in shipments to customers. And Finance and Accounting were increasingly concerned about accounts receivable problems that were attributable to errors in invoicing.

Eventually it became clear to some executives that poor master data quality was causing operational problems. But how should this be fixed? Who was responsible for master data?

A series of workshops involving key managers were organised, facilitated by a Data Management guru. At first the discussions about data were confused. Manufacturing and supply chain management had differing views and terminology about materials and products. Marketing, sales and administration had differences about channels, contracts, prospects and customers.

However, after a few cycles the guru had a series of posters telling a story about the company's data that everyone agreed on. They were surprised when he explained the posters showed conceptual data models of the primary entities of importance to the business and the beginnings of a business glossary. And that these formed the foundation for a Data Governance framework that would lead to effective rollout of Master Data Management within six months.

References

Alexopoulos, P. (2020). *Semantic modeling for data: Avoiding pitfalls and breaking dilemmas.* O'Reilly Media.

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

DAMA NL (2020). Data concepts for Data Quality Dimensions (DSC). Research paper. https://www.dama-nl.org/wp-content/uploads/2020/09/DCS-Data-Concept-System-DDQ-Research-Paper-version-1.2-d.d.-3-Sept-2020.pdf

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

Dingemans, Bert (2022). Een repository voor Meta Data Management. External Link.

ISO 11179:1999 *Information technology — Specification and standardization of data elements*

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

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

Harris, J. & Hoberman, S. (2020). *Data modeling made simple with Erwin DM.* Technics Publications LLC, New Jersey.