

Dictionary of dimensions of data quality (3DQ)

Dictionary of 60 Standardized Definitions



Colophon

Authors:

Andrew Black (Van Nederpelt & Black)

Peter van Nederpelt (Van Nederpelt & Black)

Version: See Version History

Printed of 5-9-2020 11:20

© DAMA NL Foundation. All rights reserved.

Table of contents

Tables	4
Figures	4
1. Introduction	5
1.1 Importance of data quality	5
1.2 Purpose of the dictionary	5
1.3 What can you do with the dictionary?	5
1.4 Why is this dictionary needed?	5
1.5 Principles applied in the dictionary	5
1.6 Scope	5
1.7 Audience	5
1.8 What do the terms dimensions, data, and quality mean?	6
1.9 How did the dictionary come about?	6
1.10 Release policy	6
1.11 Reading guide	6
2. The dimensions of data quality and their definitions	7
Appendix 1: Sources	16
Sources of definitions of quality dimension	16
Other sources	16
Appendix 2: Concept system dimensions of data quality	18
Appendix 3: Diagrams	21
Version history	25

Tables

Table 1: Definitions of dimensions of data quality

Table 2: Dimensions of data quality classified by data concept

Table 3: Definitions of concepts

Figures

Figure 1: Data concept system DDQ

Figure 2: Data concepts in a data model

Figure 3: Relationship between data concepts and dimensions

Figure 4: Artists impression of the real world and data world

1. Introduction

1.1 Importance of data quality

Data plays an increasingly important useful role in our society. Dependence on data for many activities and processes is increasing. Quality of data is therefore of growing importance. However, there is still no standard for the dimensions of data quality and their definitions. The purpose of this report is to take a step towards a standard.

1.2 Purpose of the dictionary

The purpose of the dictionary is to present sixty standardised definitions of dimensions of data quality. In order to manage dimensions of data quality, it is necessary to know what the meaning of these dimensions are.

1.3 What can you do with the dictionary?

The dictionary can be used as follows:

- Select relevant dimensions, i.e., dimensions that contribute to business goals.
- Establish performance indicators and associated measurement methods for the selected dimensions.
- Establish target values for these indicators.
- Measure the values of the indicators and test if they meet the target values.
- Eventually, take measures to improve the data quality.

Moreover, the dictionary can be used to promote efficient communication about dimensions of data quality by adopting the dictionary as a standard or as a reference.

1.4 Why is this dictionary needed?

The dictionary is needed because there are several sources in which dimensions of data quality are defined. However, each source presents and defines a limited number of dimensions and the quality of their definitions varies.

1.5 Principles applied in the dictionary

The following principles have been applied in compiling the definitions of the dimensions of data quality:

- The list has been made as complete as possible.
- Definitions that already exist have been used as much as possible.
- The definitions meet the requirements of ISO 704. This standard is about defining terms in general. For example, a definition should not be too long and should not contain examples.
- The definition always starts with 'the degree to which...'
- A dimension is always part of something. We call it a data concept (e.g. attribute, record or data file).
- The data concepts together form a data concept system. These data concepts are also defined.
- Dimensions of data quality can be categorized by data concept.

1.6 Scope

The scope of the report is the quality dimensions of data concepts that occur in information systems.

1.7 Audience

The report is meant for everyone who is involved in management of data quality and who wish to select their own set of dimensions. They can use the dictionary as a reference.

1.8 What do the terms dimensions, data, and quality mean?

The title of this report is Dimensions of Data Quality (DDQ). The meaning of these three words are explained below.

Dimension is defined as a measurable feature of an object (ISO 9001). The object is in this context data. The term dimension is used to make the connection to dimensions in the measurement of physical objects (e.g., length, width, height).

The term dimension in this context should not be confused with its use in the context of business intelligence where it refers to a category for summarizing or viewing data.

Data is a collective term for several concepts such as records, attributes, data values and metadata.

Quality is defined as the degree to which inherent characteristics of an object meet requirements (ISO 9001). But we reformulate this definition as: the degree to which dimensions of a data concept meet requirements.

The term 'characteristic' is replaced by 'dimension', because the latter term is more common in data land.

The term 'object' is replaced by 'data concept' to avoid confusion with objects of the real world that are represented in datasets. The term concept is, moreover, used in ISO 704, a standard for definitions.

1.9 How did the dictionary come about?

The dictionary is an initiative of the Data Quality working group of DAMA-NL. This working group drew up a research paper (Black, Van Nederpelt, 2020). Subsequently, the present report has been derived from this paper. Finally, it was submitted to the DAMA community for comment.

1.10 Release policy

The first final version of the report will be published in 2020. New versions will be compiled as needed. Proposals for changes can be made via info@dama-nl.org or the authors info@vannederpeltblack.nl.

1.11 Reading guide

Chapter 1 describes the purpose and use of the dictionary.

Chapter 2 presents the sixty dimensions of data quality and their definitions.

Appendix 1 contains references to sources for definitions.

Appendix 2 defines concepts used in the report

Appendix 3 shows diagrams of the concept system.

2. The dimensions of data quality and their definitions

This chapter demonstrates sixty dimensions of data quality. Table 1 shows the dimensions and their definitions in alphabetical order. The most common dimensions are marked blue. The dimensions are classified by data concept in Table 2.

Table 1: Definitions of dimensions of data quality

Nr	Dimension	Data Concept	Definition
1.	Ability to represent null values	Format	The degree to which a format allows null values in an attribute.
2.	Access security	Datasets	The degree to which access to datasets is restricted.
3.	Accessibility	Data	The ease with which data can be consulted or retrieved,
4.	Accuracy	Data values	The degree of closeness of data values to real values.
5.	Appropriateness	Format	The degree to which the format is suitable for use.
6.	Availability	Data	The degree to which data can be consulted or retrieved by data consumers or a process.
7.	Clarity	Metadata	The ease with which data consumers can understand the metadata.
8.	Coherence	Composition of datasets	The degree to which datasets can be combined.
9.	Comparability of populations	Data values	The degree to which data values representing two populations have the same definition and are measured in the same way.
10.	Comparability over time	Data values	The degree to which data values over time have the same definition and are measured in the same way.
11.	Completeness	Attributes	The degree to which all required attributes in the dataset are present.
12.		Records	The degree to which all required records in the dataset are present.
13.		Data files	The degree to which all required data files are present.
14.		Data values	The degree to which all required data values are present.
15.		Data values of an attribute	The degree to which all required data values of an attribute are present.

Nr	Dimension	Data Concept	Definition
16.		Metadata	The degree to which the metadata are fully described.
17.	Compliance with laws, regulations, or standards	Data	The degree to which data is in accordance with laws, regulations, or standards.
18.		Composition of datasets	The degree to which the composition of datasets is in accordance with laws, regulations, or standards.
19.	Confidentiality	Data	The degree to which disclosure of data should be restricted to authorized data consumers.
20.	Consistency	Data values	The degree to which data values of two sets of attributes <ul style="list-style-type: none"> ▪ within a record, ▪ within a data file, ▪ between data files, ▪ within a record at different points in time comply with a rule.
21.		Data values of a set of attributes of a dataset at different points in time (temporal consistency)	The degree to which the data values of a set of attributes of a dataset at different points in time comply with a rule.
22.		Data values of two sets of attributes between datasets (across datasets)	The degree to which data values of two sets of attributes between datasets comply with a rule.
23.		Data values of two sets of attributes between records (cross record)	The degree to which data values of two sets of attributes between records comply with a rule.
24.		Data values of two sets of attributes within a record (record level)	The degree to which data values of two sets of attributes within a record comply with a rule.
25.	Credibility	Data values	The degree to which data values are regarded as true and believable by data consumers.
26.	Currency	Data values	The degree to which data values are up to date.
27.	Equivalence	Attributes	The degree to which attributes stored in multiple datasets are conceptually equal.
28.	Granularity	Attributes	The degree to which a single characteristic is subdivided in attributes.

Nr	Dimension	Data Concept	Definition
29.	Granularity	Records	The degree to which objects are aggregated to records.
30.	Integrity	Data values	The degree of absence of data value loss or corruption.
31.	Interpretability	Data	The degree to which data are in an appropriate language and units of measure.
32.	Latency	Data	The period of time between the point when the data is created and the point when it is available for use.
33.	Linkability	Data files	The degree to which records of one data file can be correctly coupled with records of another data file.
34.	Metadata compliance	Data values	The degree to which the data values are in accordance with their definition, format specification and value domain.
35.	Naturalness	Composition of datasets	The degree to which the composition of datasets is aligned with the real-world objects that they represent.
36.	Objectivity	Data values	The degree to which the data values are created in an unbiased manner.
37.	Obtainability	Data	The degree to which the data can be acquired.
38.	Plausibility	Data values	The degree to which data values match knowledge of the real world.
39.	Portability	Data	The degree to which data can be installed, replaced or moved from one system to another while preserving the existing quality.
40.	Portability	Format	The degree to which a format can be applied in a wide range of situations.
41.	Precision (1)	Data values	The degree of accuracy with which data value are recorded or classified.
42.	Precision (2)	Data values	The degree to which the error in data values spreads around zero (in statistics).
43.	Punctuality	Dataset availability	The degree to which the period between the actual and target point of time of availability of a dataset is appropriate.

Nr	Dimension	Data Concept	Definition
44.	Reasonability	Data pattern	The degree to which a data pattern meets expectations.
45.	Recoverability	Datasets	The degree to which datasets are preserved in the event of incident.
46.	Redundancy	Data	The degree to which logically identical data are stored more than once.
47.	Referential integrity	Data files	The degree to which data values of the primary key of one data file and data values of the foreign key of another data file are equal.
48.	Relevance	Composition of datasets	The degree to which the composition of datasets meets the needs of the data consumer.
49.	Reliability	Initial data value	The closeness of the initial data value to the subsequent data value.
50.	Reproducibility	Dataset	The degree to which a dataset can be recreated with the same data values.
51.	Reputation	Data	The degree to which data are trusted or highly regarded in terms of their source or content.
52.	Retention period	Datasets	The period that datasets are available until they can or must be deleted.
53.	Timeliness	Dataset availability	The degree to which the period between the time of creation of the real value and the time that the dataset is available is appropriate.
54.	Traceability	Data	The degree to which data lineage is available.
55.	Uniqueness	Objects	The degree to which objects (of the real world) occur only once as a record in a data file.
56.		Records	The degree to which records occur only once in a data file.
57.	Validity	Data values	The degree to which data values comply with rules.
58.	Value	Data	The degree to which data provide advantages from their use.
59.	Variety	Data	The degree to which data are available from different data sources.
60.	Volatility	Data values	The degree to which data values change over time.

Table 2: Dimensions of data quality classified by data concept

Nr	Dimension	Data concept	Definition
		Attribute	
1.	Completeness	Attributes	The degree to which all required attributes in the dataset are present.
2.	Equivalence	Attributes	The degree to which attributes stored in multiple datasets are conceptually equal.
3.	Granularity	Attributes	The degree to which a single characteristic is subdivided in attributes.
		Data	
4.	Accessibility	Data	The ease with which data can be consulted or retrieved,
5.	Availability	Data	The degree to which data can be consulted or retrieved by data consumers or a process.
6.	Compliance with laws, regulations, or standards	Data	The degree to which data is in accordance with laws, regulations, or standards.
7.	Confidentiality	Data	The degree to which disclosure of data should be restricted to authorized data consumers.
8.	Interpretability	Data	The degree to which data are in an appropriate language and units of measure.
9.	Latency	Data	The period of time between the point when the data is created and the point when it is available for use.
10.	Portability	Data	The degree to which data can be installed, replaced or moved from one system to another while preserving the existing quality.
11.	Redundancy	Data	The degree to which logically identical data are stored more than once.
12.	Reputation	Data	The degree to which data are trusted or highly regarded in terms of their source or content.
13.	Traceability	Data	The degree to which data lineage is available.
14.	Obtainability	Data	The degree to which data can be acquired.

Nr	Dimension	Data concept	Definition
15.	Value	Data	The degree to which data provide advantages from their use.
16.	Variety	Data	The degree to which data are available from different data sources.
		Data files	
17.	Completeness	Data files	The degree to which all required data files are present.
18.	Linkability	Data files	The degree to which records of one data file can be correctly coupled with records of another data file.
19.	Referential integrity	Data files	The degree to which data values of the primary key of one data file and data values of the foreign key of another data file are equal.
		Data pattern	
20.	Reasonability	Data pattern	The degree to which a data pattern meets expectations.
		Dataset	
21.	Access security	Datasets	The degree to which access to datasets is restricted.
22.	Recoverability	Datasets	The degree to which datasets are preserved in the event of incident.
23.	Reproducibility	Dataset	The degree to which a dataset can be recreated with the same data values.
24.	Retention period	Datasets	The period that datasets are available until they can or must be deleted.
		Dataset availability	
25.	Punctuality	Dataset availability	The degree to which the period between the actual and target point of time of availability of a dataset is appropriate.
26.	Timeliness	Dataset availability	The degree to which the period between the time of creation of the real value and the time that the dataset is available is appropriate.
		Dataset composition	
27.	Coherence	Composition of datasets	The degree to which datasets can be combined.
28.	Compliance to laws, regulations, or standards	Composition of datasets	The degree to which the composition of datasets is in accordance with laws, regulations or standards.

Nr	Dimension	Data concept	Definition
29.	Naturalness	Composition of datasets	The degree to which the composition of datasets is aligned with the real-world objects that they represent.
30.	Relevance	Composition of datasets	The degree to which the composition of datasets meets the needs of the data consumer.
		Data value	
31.	Accuracy	Data values	The degree of closeness of data values to real values.
32.	Comparability of populations	Data values	The degree to which data values representing two populations have the same definition and are measured in the same way.
33.	Comparability over time	Data values	The degree to which data values over time have the same definition and are measured in the same way.
34.	Completeness	Data values	The degree to which data values are present.
35.	Consistency	Data values	The degree to which data values of two sets of attributes <ul style="list-style-type: none"> ▪ within a record, ▪ within a data file, ▪ between data files, ▪ within a record at different points in time comply with a rule.
36.	Credibility	Data values	The degree to which data values are regarded as true and believable by data consumers.
37.	Currency	Data values	The degree to which data values are up to date.
38.	Integrity	Data values	The degree of absence of data value loss or corruption.
39.	Metadata compliance	Data values	The degree to which the data values are in accordance with their definition, format specification and value domain.
40.	Objectivity	Data values	The degree to which the data values are created in an unbiased manner.
41.	Plausibility	Data values	The degree to which data values match knowledge of the real world.
42.	Precision (1)	Data values	The degree of accuracy with which data value are recorded or classified.

Nr	Dimension	Data concept	Definition
43.	Precision (2)	Data values	The degree to which the error in data values spreads around zero (in statistics).
44.	Validity	Data values	The degree to which data values comply with rules.
45.	Volatility	Data values	The degree to which data values change over time.
46.	Completeness	Data values of an attribute	The degree to which all required data values of an attribute are present.
47.	Consistency	Data values of a set of attributes of a dataset at different points in time (temporal consistency)	The degree to which the data values of a set of attributes of a dataset at different points in time comply with a rule.
48.	Consistency	Data values of two sets of attributes between datasets (across datasets)	The degree to which data values of two sets of attributes between datasets comply with a rule.
49.	Consistency	Data values of two sets of attributes between records (cross record)	The degree to which data values of two sets of attributes between records comply with a rule.
50.	Consistency	Data values of two sets of attributes within a record (record level)	The degree to which data values of two sets of attributes within a record comply with a rule.
		Initial data value	
51.	Reliability	Initial data value	The closeness of the initial data value to the subsequent data value.
		Format	
52.	Ability to represent null values	Format	The degree to which a format allows null values in an attribute.
53.	Appropriateness	Format	The degree to which the format is suitable for use.
54.	Portability	Format	The degree to which a format can be applied in a wide range of situations.
		Metadata	
55.	Clarity	Metadata	The ease by which data consumers can understand the metadata of a dataset.
56.	Completeness	Metadata	The degree to which the metadata are fully described.
		Object	
57.	Uniqueness	Objects	The degree to which objects (of the real world) occur only once as a record in a dataset.
		Record	

Nr	Dimension	Data concept	Definition
58.	Completeness	Records	The degree to which all required records in the dataset are present.
59.	Granularity	Records	The degree to which objects are aggregated to records.
60.	Uniqueness	Records	The degree to which records occur only once in a dataset.

Appendix 1: Sources

Sources of definitions of quality dimension

- Brackett, M. H. (2012). *Data Resource Integration: Understanding and Resolving a Disparate Data Resource* (1st ed.). Bradley Beach, NJ: Technics Publications, LLC.
- CDDQ. (n.d.). List of Conformed Dimensions of Data Quality. Retrieved from <https://dimensionsofdataquality.com/alldimensions>
- Daas, P.J.H., Ossen, S.J.L., & Tennekes, M. (2010). *Determination of Administrative Data Quality: Recent results and new developments*. Retrieved from Q2010 website: https://q2010.stat.fi/media/presentations/special-session-34/daas_ossen_tennekes_q2010_paper_session34_piet_daas_paper.pdf
- DAMA (2017). DAMA-DMBOK. *Data Management Body of Knowledge*. 2nd Edition. Technics Publications LLC. August 2017.
- DAMA-UK (2013). *The six primary dimensions for data quality assessment*. October 2013.
- Earley, S. (2011). *The DAMA Dictionary of Data Management* (2nded.). NJ: Technics Publications LLC.
- English, L. P. (1999). *Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*. Hoboken, NJ: Wiley.
- Eurostat. (2015). *ESS Handbook for Quality Reports*. (2015). Brussels, Belgium: Eurostat.
- Everest. (2010).
- Fisher, Craig; Lauria, Eitel; Chengalur-Smith, Shobha; Wang, Richard (2011). *Introduction to Information Quality*. Bloomington: AuthorHouse.
- ISO 25012. (n.d.). Retrieved from <https://iso25000.com/index.php/en/iso-25000-standards/iso-25012>
- Jayawardene, J., Sadiq, S., & Indulska, M. (2015). An analysis of data quality dimensions. *Computer Science*. Retrieved from <https://pdfs.semanticscholar.org/3d9a/c49f03b3e4bebae7c0e7eb20d7fde7222a9c.pdf>
- Nederpelt, P.W.M. van (2009). *Checklist Quality of Statistical Output*. Den Haag/Heerlen: Centraal Bureau voor de Statistiek.
- Redman, T. C. (1996). *Data Quality for the Information Age*. Artech House on Demand.
- Wang, R.Y. and Strong, D. (1996). *Beyond Accuracy: What data quality means to Data Consumers*. *Journal of Management Information Systems*, 1996. 12(4): p. 5 – 34.

Other sources

- Black, A., Nederpelt, P. van. (2020). *Data concept system for Data Quality Dimensions*. *Research paper*. DAMA-NL.
- Black, A., Nederpelt, P. van. (2020). *Dimensions of Data Quality Dimensions*. *Research paper*. DAMA-NL.
- Brackett, M. H. (2012). *Data Resource Integration: Understanding and Resolving a Disparate Data Resource* (1st ed.). Bradley Beach, NJ: Technics Publications, LLC.
- Business Directory. (2020, 1 mei). Geraadpleegd op 2 mei 2020, van <http://www.businessdictionary.com/definition/conceptual-framework.html>
- Cambridge English dictionary. (2020, August 1). *Meanings & definitions*. Cambridge Dictionary | English Dictionary, Translations & Thesaurus. <https://dictionary.cambridge.org/dictionary/english>
- DAMA (2017). DAMA-DMBOK. *Data Management Body of Knowledge*. 2nd Edition. Technics Publications Llc. August 2017.
- Earley, S. (2011). *The DAMA Dictionary of Data Management* (2nd ed.). NJ: Technics Publications Llc.

- Hay, D. (2013). *Data model patterns: Conventions of thought*. Addison-Wesley.
- Humbley, J., Budin, G., & Laurén, C. (2018). *Languages for Special Purposes*. Berlin, Germany: De Gruyter.
- Investopedia. (2015, January 7). *Patterns vs. trends: What's the difference?* <https://www.investopedia.com/ask/answers/010715/what-are-differences-between-patterns-and-trends.asp>
- ISO 1087 (2019). *Terminology work and terminology science — Vocabulary*. Vernier, Switzerland: ISO.
- ISO 11179-1 (1999). *Information technology — Specification and standardization of data elements — Part 1: Framework for the specification and standardization of data elements*. ISO
- ISO 21961 (2003). *Space data and information transfer systems – Data entity dictionary specification language (DEDSL) – Abstract syntax*. Retrieved from <https://www.iso.org/obp/ui#iso:std:iso:21961:ed-1:v1:en>.
- ISO 704 (2009). *Technical Committee ISO/TC 37; Terminology and other language and content resources. Subcommittee SC 1; Principles and Methods*. (2009). *Terminology Work – Principles and Methods*. Vernier, Switzerland: ISO.
- ISO 9000:2015 (2015). *Quality management systems – Fundamentals and vocabulary*. Delft: NNI.
- ISO 9001:2015 (2015). *Quality management systems – Requirements*. Delft: NNI.
- Jonker R. (2020, May 6). *Terminologie*. Retrieved from <https://labyrinth.rienkjonker.nl/lexicon/terminologie>
- Lexico Dictionaries (2020, August 1). *Definitions, meanings, synonyms, and grammar by Oxford dictionary on Lexico.com*. Lexico Dictionaries | English. <https://www.lexico.com>
- McGilvray, D. (2008). *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*. Morgan Kaufman Publishers.
- Merriam-Webster. (2020, May 7). *Dictionary by Merriam-Webster: America's most-trusted online dictionary*. Retrieved from <https://www.merriam-webster.com/dictionary>
- Nederpelt, P.W.M. van (2012). *Object-oriented Quality and Risk Management. A practical method for quality and risk management*. New York/Alphen den Rijn: Lulu/MicroData.
- Oxford. (2020, May 7). *Oxford learner's dictionaries | Find definitions, translations, and grammar explanations at Oxford learner's dictionaries*. Retrieved from <https://www.oxfordlearnersdictionaries.com>
- Regoniel, P. (2015, January 5). *Conceptual framework: A step-by-step guide on how to make one*. Retrieved from <https://simplyeducate.me/2015/01/05/conceptual-framework-guide/>
- UN/Edifact glossary. (n.d.). Retrieved from https://www.unece.org/trade/untdid/texts/d300_s.htm

Appendix 2: Concept system dimensions of data quality

In this appendix concepts are defined that are relevant in this report. See Table 3. These concepts are taken from the report Concept System of Dimensions of Data Quality (Black, Van Nederpelt, 2020).

A distinction is made between data concepts in de real world (purple) and data world (yellow). Other concepts (white) are more general.

Each word that appears in **bold** in the definition of a concept is a concept defined elsewhere in Table 3. This way the coherence between the concepts are made visible.

Table 3: Definitions of concepts

Concept	Definition	Source	Relationships with other concepts
Attribute	A characteristic of an entity type about which the organisation wishes to hold information.	-	Distinguishes entity type Is specified by its name, definition, classification and format .
Characteristic	Distinguishing feature	ISO 9000	-
Composition of a dataset	The way in which a dataset is made up.	-	
Concept	Unit of knowledge created by a unique combination of characteristics	ISO 1087	-
Concept system	A set of concepts structured according to the relations among them.	ISO 704	-
Data	A representation of facts, concepts, or instructions in a formalized manner, suitable for communication, interpretation, or processing by humans or by automatic means. (ISO 2382-4).	In: ISO 11179	-
Data category	A classification of data according to the purpose for which it is used.	-	-
Data concept	A concept related to data	-	Has associated dimensions
Data file	Data stored on a computer as one unit with one name.	Cambridge 2020	Is part of a dataset .
Data item	One occurrence of an attribute	-	Contains data value
Data lineage	Metadata that identifies the sources of data and the transformations through which it has passed up to the point of consumption.	-	-

Concept	Definition	Source	Relationships with other concepts
Data pattern	A series of data that repeats in a recognizable way.	Investopedia	-
Data value	The value of a data item .	-	Is contained in data item Forms part of record Is within value domain Represents a property of an object
Dataset	Any organized collection of data .	Early 2011	Is composed of data files
Dataset availability	The degree to which a dataset can be consulted or retrieved by data consumers or processes.	-	Is a characteristic of a dataset .
Dataset composition	The way in which a dataset is made up.	-	-
Definition	Representation of a concept by an expression that describes it and differentiates it from related concepts	ISO 1087	-
Dimension	Measurable characteristic .	DAMA 2017	Is associated with a data concept .
Entity type	A thing of significance about which the organisation wishes to hold information	Hay 2013	Is distinguished by attributes Describes object
Initial data value	A provisional data value that will be updated by a more accurate value.	-	Is a specification of a data value .
Format	A combination of datatype, unit of measure and character set.	-	Is part of the specification of an attribute .
Metadata	Data that defines and describes other data .	ISO 11179	Is a an instance of data category .
Master Data	Data held by an organization which describe object types that it needs to reference in order to perform its transactions.	-	Is an instance of data category .
Object	Anything perceivable or conceivable.	ISO 9000	Is described by entity type Is characterised by properties Is represented by records

Concept	Definition	Source	Relationships with other concepts
Property	A feature of an object .	ISO 1087	Characterises object Is recorded by data value Actually, has real value
Register	A dataset designated by the government in which vital data about citizens, residents, companies, institutions, vehicles, topography, buildings, and addresses can be centrally maintained.	-	Is an instance of data category .
Statistical output	Output from a statistical process.	-	Is an instance of data category .
Transactional data	Data that describes an event that takes place as an organization conducts its business.	-	Is an instance of data category .
Real value	The real-life value of a property of an object .	-	Expresses an instance of a property .
Reference data	Data used to categorize other data.	-	Is an instance of data category .
Record	A logically related set of data values that represent a (real-world) object	-	Forms part of data file Is composed of data values
Value domain	A set of permissible values of an attribute .	-	Includes data value

Appendix 3: Diagrams

The data concept system relevant for dimensions of data quality (DDQ) is shown in Figure 1.

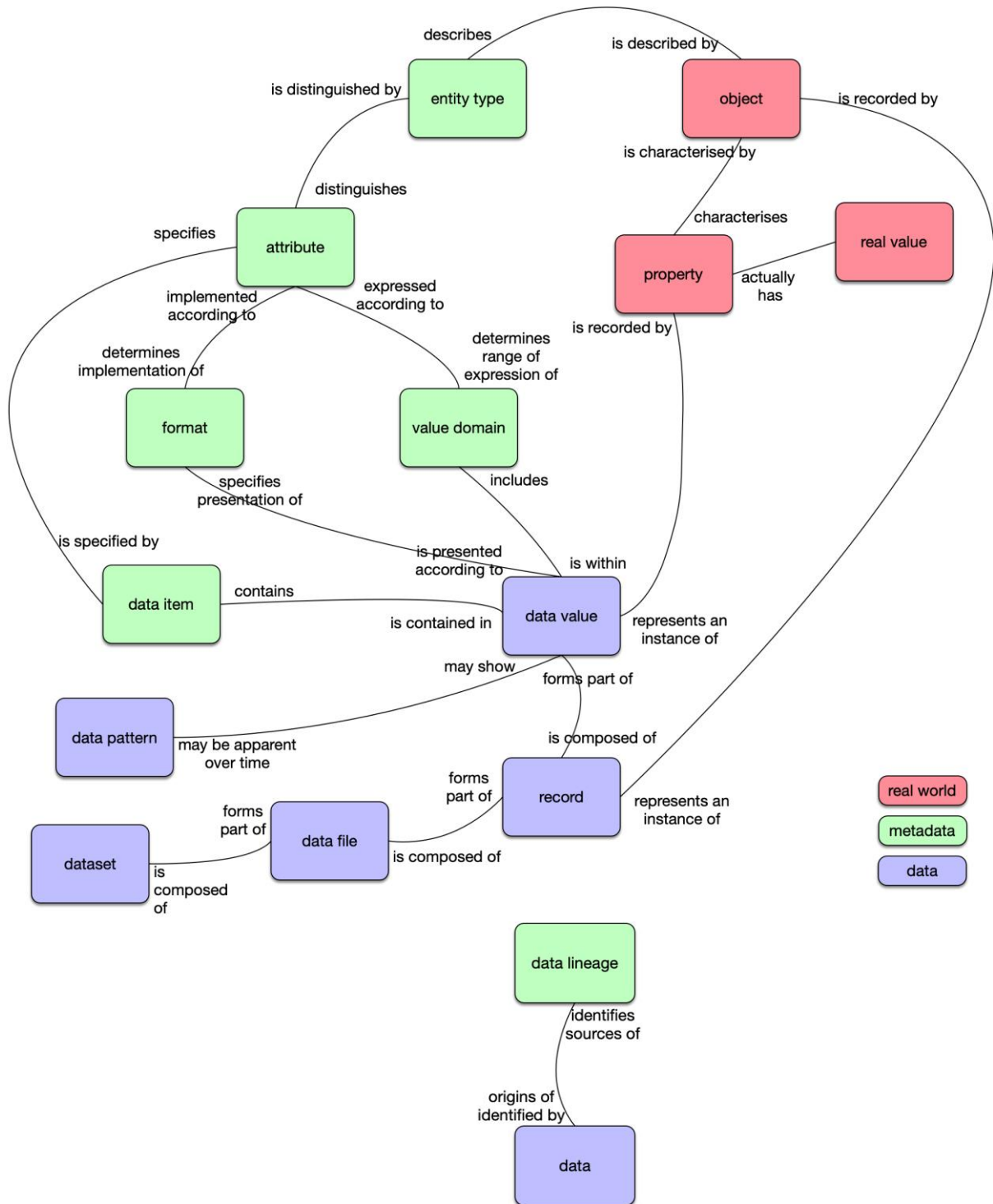


Figure 1: Data concept system DDQ

A set of data concepts in de data world is shown in a data model in Figure 2.

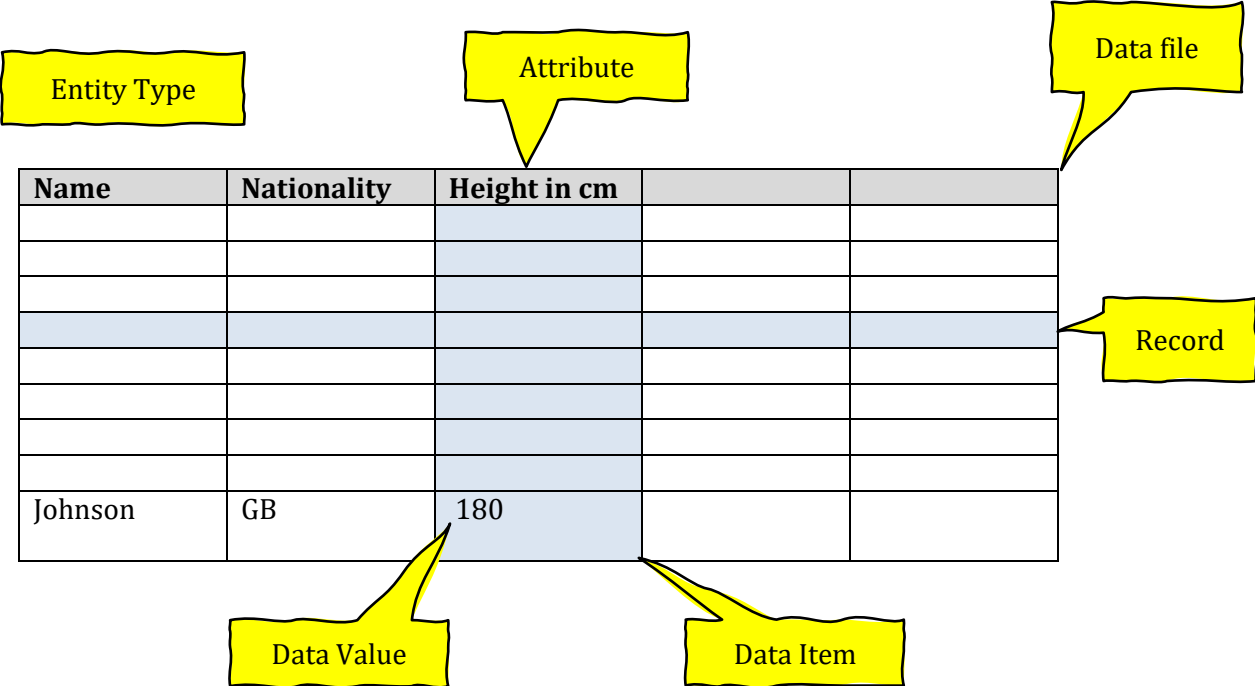


Figure 2: Data concepts in a data model

Figure 3 shows that a dimension is associated with a data concept. The definition of a dimension of data quality is formed by the combination of a dimension and a data concept.



Figure 3: Relationship between data concepts and dimensions

Figure 4 is an artist's impression of the real world and data world.

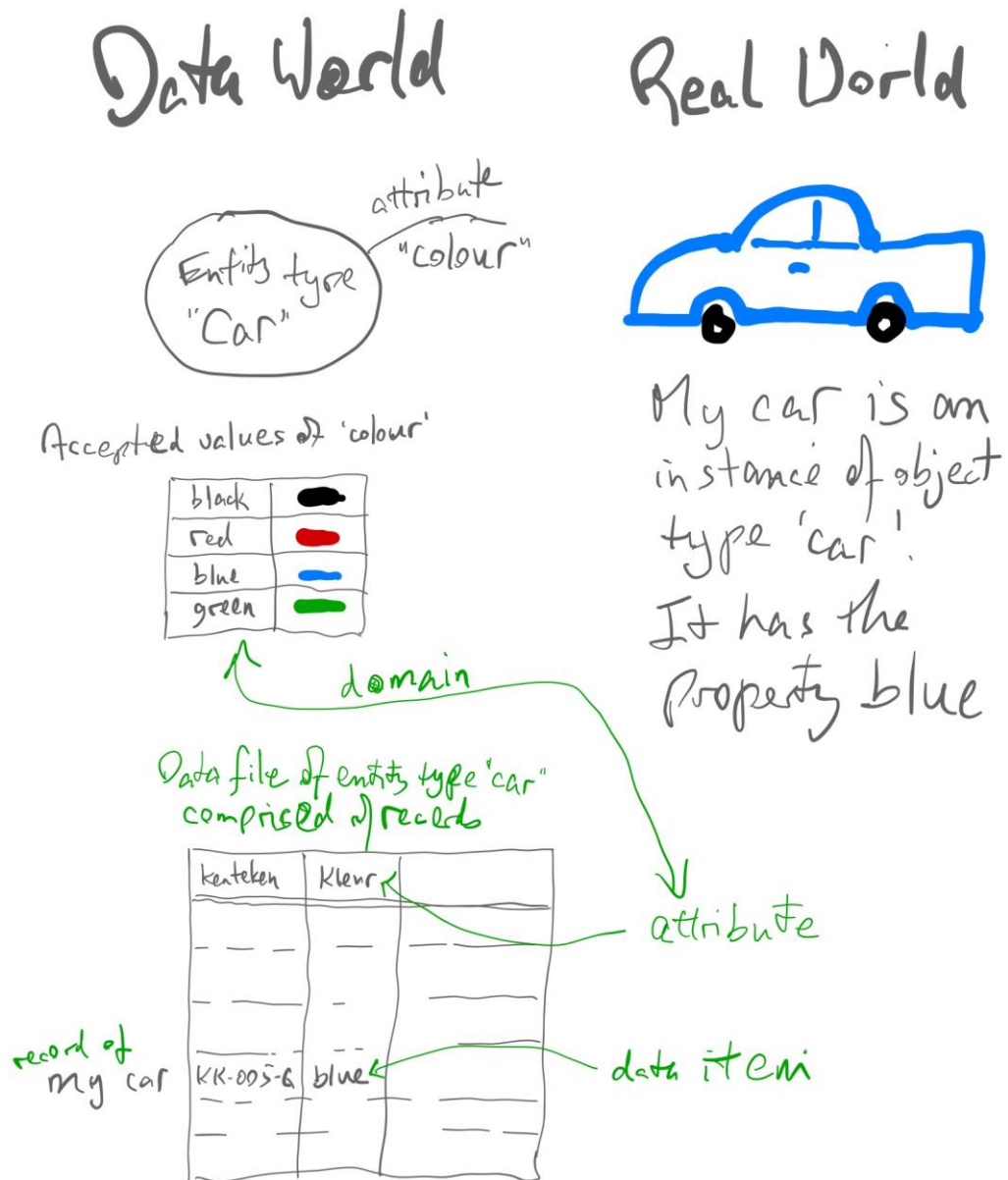


Figure 4: Artist's impression of the real world and data world

Version history

Version	Date	Description of the modification	Author
1.0.p1	5 August 2020	First draft	Peter
1.0.p2	10 August 2020	Comments and minor edits	Andrew
1.0	20 August 2020	Edits and comments processed. Three data concepts added. Final version.	Peter
1.1	3 Sept 2020	Data categories and one source added.	Peter

Active distribution per version	
Version	Distribution
1.0.p1-p2	Dropbox
1.0	Dropbox
1.1	Dropbox. Website DAMA-NL