INTERNATIONAL STANDARD

ISO 8000-100

First edition
2016-10-01

Data quality —

Part 100: Master data: Exchange of characteristic data: Overview

Qualité des données —

Partie 100: Données permanentes: Échange des données caractéristiques: Aperçu général

Reference number
ISO 8000-100:2016(E)

COPYRIGHT PROTECTED DOCUMENT

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is Technical Committee ISO/TC 184, *Automation systems and integration*, Subcommittee SC 4, *Industrial data*.

This first edition of ISO 8000-100 cancels and replaces ISO/TS 8000-100:2009, which has been technically revised.

ISO 8000 is organized as a series of parts, each published separately. The structure of ISO 8000 is described in ISO/TS 8000-1.

Each part of ISO 8000 is a member of one of the following series: general data quality, master data quality, transactional data quality, and product data quality. This part of ISO 8000 is a member of the master data quality series.

A list of all parts in the ISO 8000 series can be found on the ISO website.

Introduction

The ability to create, collect, store, maintain, transfer, process and present data to support business processes in a timely and cost effective manner requires both an understanding of the characteristics of the data that determine its quality, and an ability to measure, manage and report on data quality.

ISO 8000 defines characteristics that can be tested by any organization in the data supply chain to objectively determine conformance of the data to ISO 8000.

ISO 8000 provides frameworks for improving data quality for specific kinds of data. The frameworks can be used independently or in conjunction with quality management systems.

ISO 8000 covers industrial data quality characteristics throughout the product life cycle from conception to disposal. ISO 8000 addresses specific kinds of data including, but not limited to, master data, transaction data, and product data.

The master data quality series of parts of ISO 8000 addresses the quality of master data. This part of ISO 8000 is an introduction to the series. It contains an introduction to master data, a data architecture, a high-level data model, and an overview of the remaining parts of the series.

Annex A contains an identifier that unambiguously identifies this part of ISO 8000 in an open information system.

Annex B describes different categories of items and their identifiers.

Data quality —

Part 100:
Master data: Exchange of characteristic data: Overview

1 Scope

This part of ISO 8000 contains an overview of the master data quality series of parts of ISO 8000, which addresses master data quality.

The following are within the scope of the master data quality series of parts of ISO 8000:

— master data-specific aspects of quality management systems;

— master data quality metrics.

The approach of the master data quality series of parts of ISO 8000 is to address data quality:

— from the bottom up, i.e. from the smallest meaningful element, the property value;

— at the interface of master data management systems, not within the systems.

The master data quality series of parts of ISO 8000 contains requirements that can be checked by computer for the exchange, between organizations and systems, of master data that consists of characteristic data. These parts address the quality of property values that are exchanged within master data messages.

This part of ISO 8000 describes fundamentals of master data quality and specifies requirements on both data and organizations to enable master data quality.

The following are within the scope of this part of ISO 8000:

— specification of the scope of the master data quality series of parts of ISO 8000;

— introduction to master data;

— description of the data architecture;

— overview of the content of the other parts of the series.

The following are outside the scope of this part of ISO 8000:

— aspects of data quality that apply to all data regardless of whether they are master data;

— aspects of data quality that apply to data that are not master data.

EXAMPLE Transaction data are not considered to be master data.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 8000-2, *Data quality — Part 2: Vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 8000-2 apply.

4 Abbreviated terms

MDR master data record

UML Unified Modeling Language

5 Master data

Within an organization, master data is used to identify and describe things that are significant to the organization.

NOTE 1 In cataloguing applications, master data are used to describe things called “items”.

Figure 1 depicts a taxonomy of data, showing where master data fits.

NOTE 2 Figure 1 is not intended to be a complete taxonomy of data; it is only intended to show the context of master data.

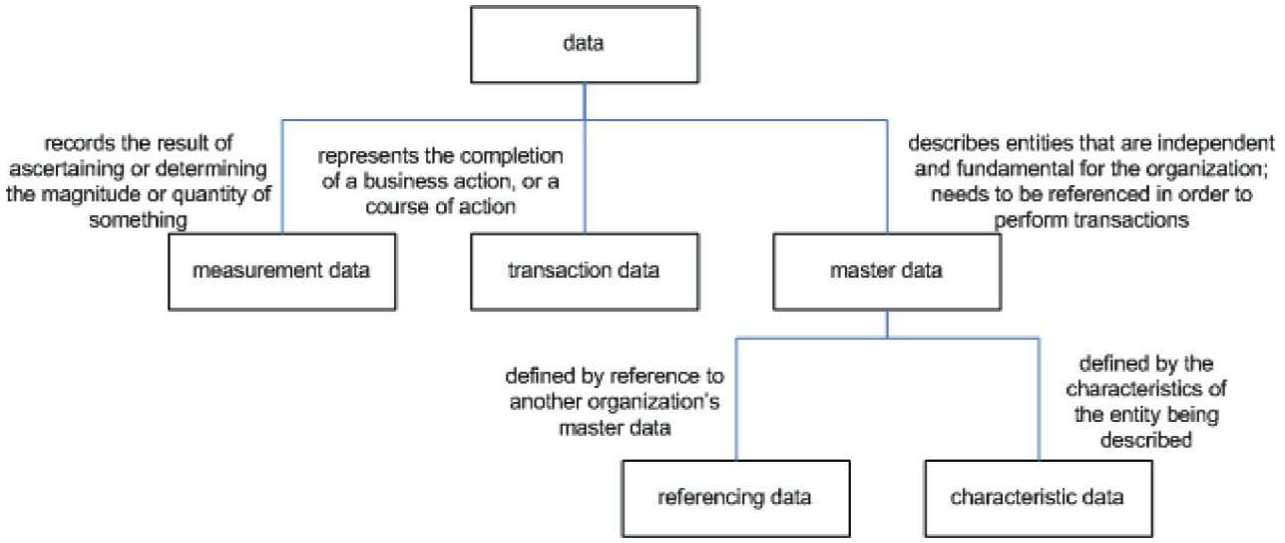

Figure 1 — Taxonomy of data (for master data)

Master data is typically referenced in business transactions through an identifier. The identifier is commonly a reference both to the thing itself and to a master data record (MDR) that describes the thing. The MDR is commonly held in a central repository.

EXAMPLE 1 It is common for the central repository of MDRs for an organization to be the organization’s enterprise resource planning (ERP) or master data management (MDM) system.

NOTE 3 What is logically a single MDR can be represented by several physical records in a software system.

EXAMPLE 2 In a relational database implementation, a master data record could consist of rows from several different tables.

NOTE 4 A MDR that describes something can be identified via a reference using its identifier. Something can be described by characteristic data, represented by property values. Additionally, something can be described by descriptive strings or definitions.

Identifying references are designed to be used as references to master data held by others.

EXAMPLE 3 A corporate tax identifier, an individual's national insurance number, and a part number assigned by a manufacturer to an item of production are all examples of identifying references.

In order for an identifying reference to be meaningful, it shall be associated with a system of identification.

EXAMPLE 4 The organization that issued the identifier can be specified by the metadata, as is common in tax identifiers, but a part number is meaningless if the manufacturer that issued it is not known.

A description can be computer interpretable characteristic data, which is typically represented as property values, or human readable text. Some properties are differentiating. Because of the ease with which they can be processed, numerical or controlled values are most often used as differentiating.

One of the key aspects of managing master data quality is managing duplication. A consistent approach to managing and eliminating inappropriate duplication is a critical part of master data management.

A characteristic that is considered differentiating by one organization could be considered non-differentiating by another organization.

EXAMPLE 5 A manufacturer would have a different master data record for each of its items of production. When, from a buyer's perspective, several items of production (produced by the same manufacturer or different manufacturers) share the same characteristics of fit, form and function, the buyer may group under a single item of supply and assign a "stock number" as the identifying reference for the item of supply. In grouping several items of production as a single item of supply, the buyer is making a decision to consider as non-differentiating one or more characteristics that the manufacturer(s) consider differentiating.

A characteristic that is considered differentiating by one function within an organization may be considered non-differentiating by another function within the same organization.

Master data is not necessarily static. Also, the number of characteristics needed to describe something will vary by business function. As the number of differentiating characteristics various, MDRs may have to be differentiated when characteristics are added or changed to differentiating. MDRs may become duplicates when characteristics are removed or changed to be non-differentiating.

Examples of master data include:

— vendor master: This typically describes a vendor in term of its location and legal status. Much of the mandatory data in a vendor master is prescribed by law as it is a common requirement for a company to be able to identify all entities to which it has transferred funds.

— customer master: This typically describes a customer in terms of a trading entity. At a minimum it will include the contact information necessary to transmit invoices and may contain confidential information such as credit card information.

  NOTE 5 If personal data are maintained in a customer master, it can be subject to data protection legislation.

— item or material master: These masters typically describe tangible items that are tracked, inventoried or regularly purchased. While they are often restricted to items purchased under contract such as production materials they can also be used to improve the quality of spend analysis associated with maintenance, repair and operations (MRO) purchases. Material masters are also commonly used to support bills of materials (BOM) or to in design where they may be referred to as common parts catalogue or a preferred part list. A variation of the material master is an illustrated parts catalogue (IPC) or a spare parts list.

— item of supply concept: These masters include a reference to an item or material master, plus packaging and quantity information;

— service, procedure or process master: These masters are still relatively rare except in the health care and vehicle repair industries where automated billing for services or insurance reimbursement is common. Typically a service is best described as a procedure or a process.

EXAMPLE 6 The American Medical Association's Current Procedural Terminology-4 (CPT-4) codes is an example of a procedure master.

— asset master: These masters are commonly used to track items whose purchase price is over a preset monetary value, or whose cost is depreciated over several years. Assets are commonly associated with a unique identifier (serial number) and often associated with movable items where date (time occasionally) and location need to be verified and reported. Correct modelling of an asset master is important to be able to track not only the location and value of the asset over time but also the maintenance and repair activity. A typical problem with asset management is changing specifications over the asset's life span. Deciding at what stage an asset has been so modified as to require the creation of a newly described asset is often a challenging issue.

— location master: Other than delivery services it is rare to see a separate location master, yet separating out the location master from customer and vendor masters typically leads to improved data quality. The data model for a location master is basically simple as in theory it describes a physical location where global positioning coordinates provide the absolute reference. In practice there may need to include other delivery instructions such as a postal address.

— point of contact, employee or human resource master: These masters typically describe an individual. Commonly they include information related to the relationship with the employing organization but these are better treated as transaction data as opposed to master data. They often contain confidential information.

NOTE 6 The data contained in these masters can be regulated and subject to data protection legislation.

6 Data architecture for master data

This clause contains a high-level architecture of master data. This architecture could have applicability beyond master data.

NOTE A more general architecture is intended to be incorporated into ISO/TS 8000-1.

Figure 2 shows the data architecture for master data.

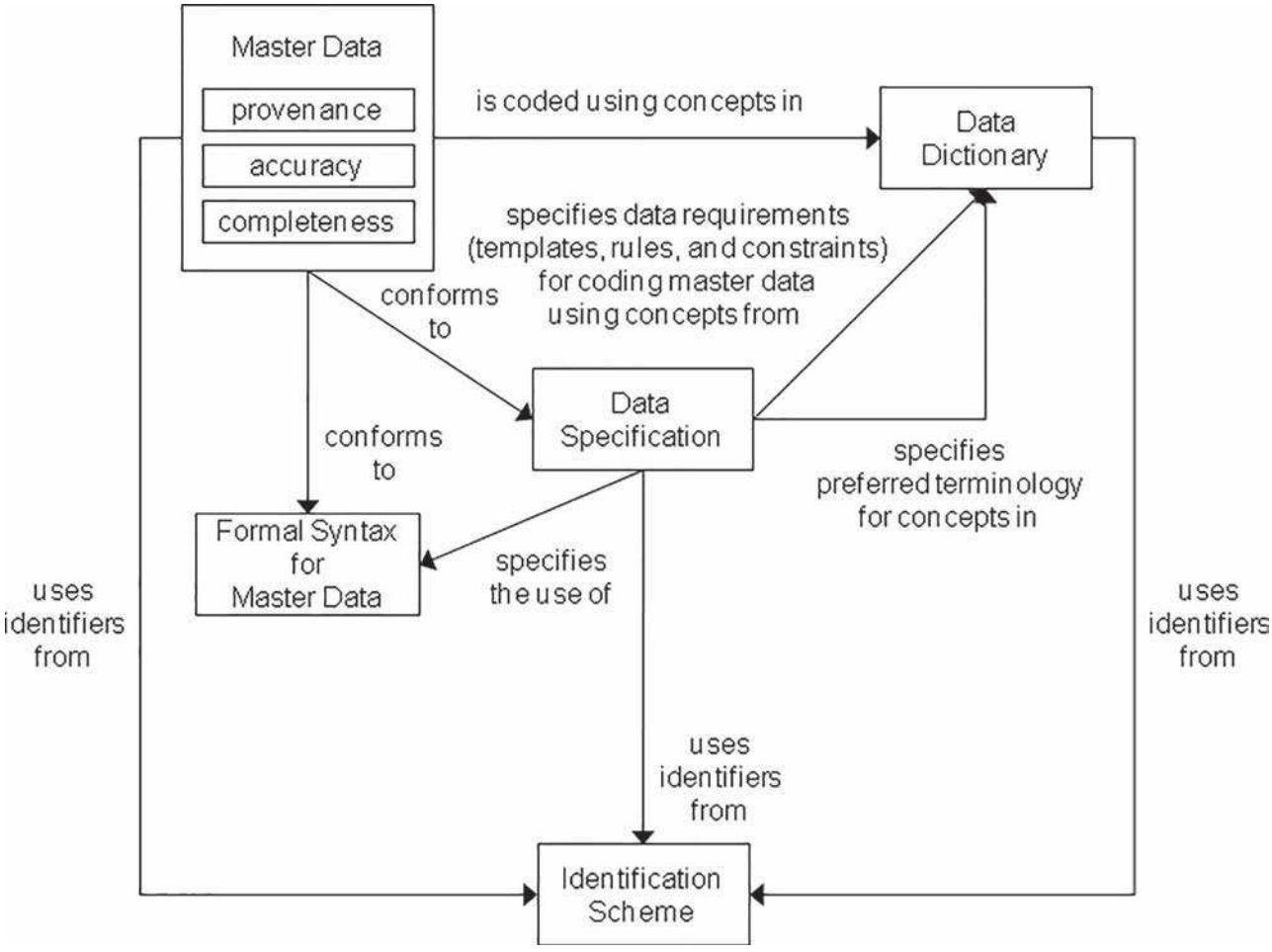

Figure 2 — Data architecture for master data

Master data includes information about data provenance, data accuracy, and data completeness.

Master data is coded using concepts in a data dictionary.

Master data conforms to a data specification.

Master data conforms to a formal syntax.

A data specification specifies data requirements for coding master data using concepts from a data dictionary.

A data specification specifies preferred terminology for concepts in a data dictionary.

A data specification specifies the use of a formal syntax.

Master data, data specifications, and data dictionaries use identifiers from an identification scheme.

7 High-level data model

7.1 General

Clause 7 contains a high-level conceptual data model of master data. The purpose of this data model is to provide a bird's-eye view of how the key entities in the data models in ISO 8000-110, ISO 8000-120, ISO 8000-130, and ISO 8000-140 fit together. This data model could have applicability beyond master data.

NOTE A more general conceptual data model is intended to be incorporated into ISO/TS 8000-1.

7.2 Diagram

The UML class diagram for the high-level data model model is given in Figure 3.

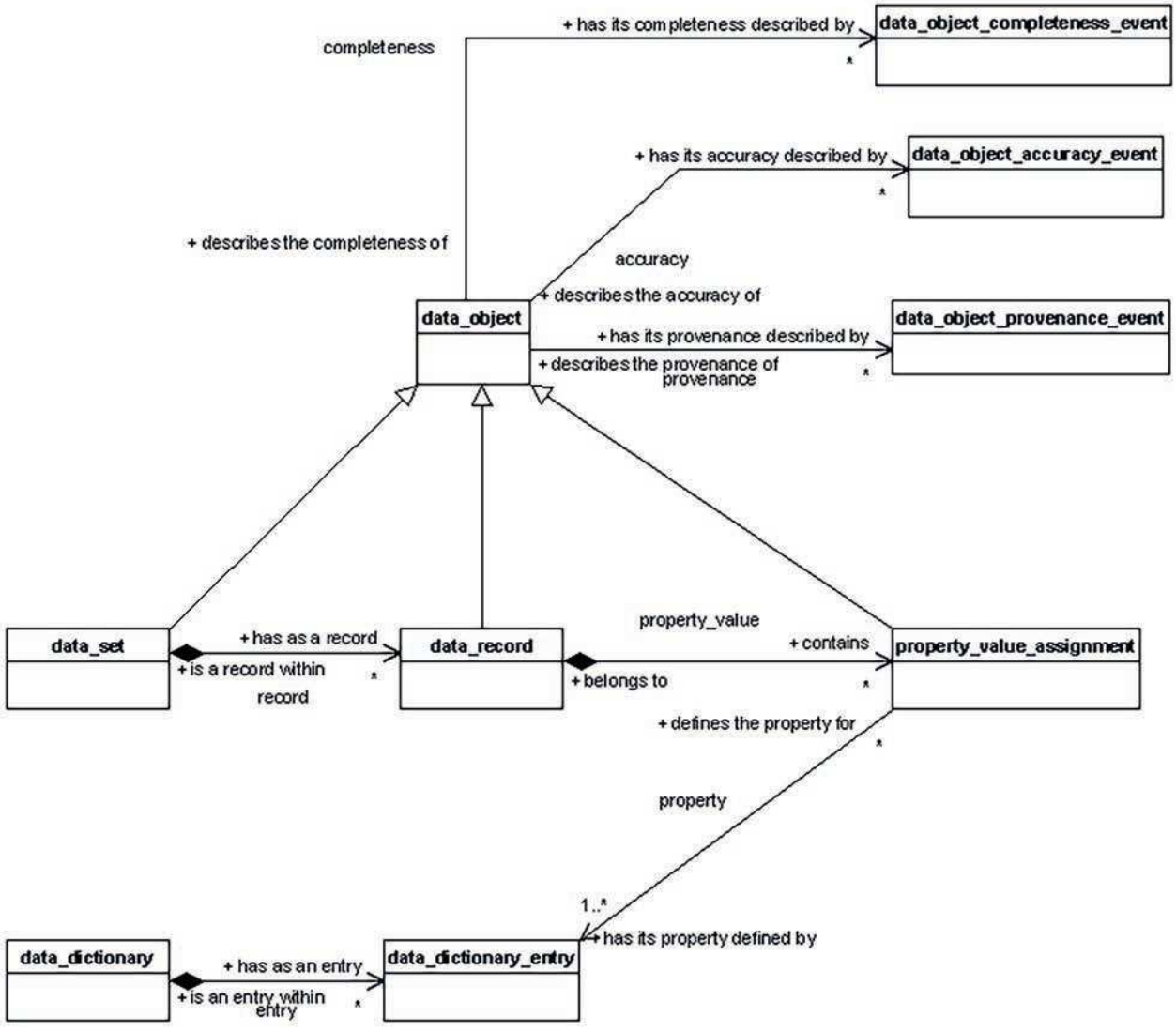

Figure 3 — UML class diagram for high-level data model

7.3 Entities

7.3.1 data_dictionary

A data_dictionary is a collection of data_dictionary_entry objects that allows lookup by entity identifier.

Attribute definitions:

entry: the data_dictionary_entry that makes up the data_dictionary.

Assertions:

Each data_dictionary has as an entry zero, one, or many data_dictionary_entry objects. Each data_dictionary_entry is an entry within exactly one data_dictionary.

7.3.2 data_dictionary_entry

A data_dictionary_entry is a description of an entity containing, at a minimum, an unambiguous identifier, a term, and a definition.

Assertions:

Each data_dictionary_entry is an entry within exactly one data_dictionary. Each data_dictionary has as an entry zero, one, or many data_dictionary_entry objects.

Each data_dictionary_entry defines the property for zero, one, or many property_value_assignment objects. Each property_value_assignment has its property defined by one or many data_dictionary_entry objects.

7.3.3 data_record

A data_record is a data_object that is a set of property_value_assignment objects.

NOTE The property can be implicit from the order of the property value pairs.

Attribute definitions:

property_value: the property_value_assignment that makes up the data_record.

Assertions:

Each data_record contains zero, one, or many property_value_assignment objects. Each property_value_assignment belongs to exactly one data_record.

Each data_record is a record within exactly one data_set. Each data_set has as a record zero, one, or many data_record objects.

7.3.4 data_set

A data_set is a data_object that is a set of data_record objects, which may be ordered or partially ordered.

Attribute definitions:

record: the data_record that makes up the data_set.

Assertions:

Each data_set has as a record zero, one, or many data_record objects. Each data_record is a record within exactly one data_set.

7.3.5 data_object

A data_object is anything that is used to signify something else.

Attribute definitions:

accuracy: the data_object_accuracy_event that provides information on the accuracy of the data_object.

completeness: the data_object_completeness_event that provides information on the completeness of the data_object.

provenance: the data_object_provenance_event that provides information on the provenance of the data_object.

Assertions:

Each data_object has its provenance described by zero, one, or many data_object_provenance_event objects. Each data_object_provenance_event describes the provenance of exactly one data_object.

Each data_object has its accuracy described by zero, one, or many data_object_accuracy_event objects. Each data_object_accuracy_event describes the accuracy of exactly one data_object.

Each data_object has its completeness described by zero, one, or many data_object_completeness_event objects. Each data_object_completeness_event describes the completeness of exactly one data_object.

7.3.6 data_object_accuracy_event

A data_object_accuracy_event is an event for which data accuracy information is recorded.

Assertions:

Each data_object_accuracy_event describes the accuracy of exactly one data_object. Each data_object has its accuracy described by zero, one, or many data_object_accuracy_event objects.

7.3.7 data_object_completeness_event

A data_object_completeness_event is an event for which data completeness information is recorded.

Assertions:

Each data_object_completeness_event describes the completeness of exactly one data_object. Each data_object has its completeness described by zero, one, or many data_object_completeness_event objects.

7.3.8 data_object_provenance_event

A data_object_provenance_event is an event for which data provenance information is recorded.

Assertions:

Each data_object_provenance_event describes the provenance of exactly one data_object. Each data_object has its provenance described by zero, one, or many data_object_provenance_event objects.

7.3.9 property_value_assignment

A property_value_assignment is a data_object that is a pair of a value and an identifier to a property defined in a data dictionary.

Attribute definitions:

ID: the string that unambiguously identifies the property_value_assignment within the organization that created it.

NOTE 1 The identifier need only be unique or meaningful within the organization that created the property_value_assignment.

NOTE 2 The format of the identifier is not specified in this part of ISO 8000.

property: the data_dictionary_entry that defines the property to which a value is being assigned.

Assertions:

Each property_value_assignment belongs to exactly one data_record. Each data_record contains zero, one, or many property_value_assignment objects.

Each property_value_assignment has its property defined by one or many data_dictionary_entry objects. Each data_dictionary_entry defines the property for zero, one, or many property_value_assignment objects.

8 Overview of the master data quality series of parts of ISO 8000

The master data quality series of parts of ISO 8000 consists of the following parts.

This part of ISO 8000 includes the following:

— scope of the master data quality series of parts of ISO 8000;

— introduction to master data;

— data architecture of the master data quality series of parts of ISO 8000;

— overview of the content of the other parts of the series.

ISO 8000-110, ISO 8000-120, ISO 8000-130 and ISO 8000-140 contain requirements that can be checked by computer for the exchange, between organizations and systems, of master data that consists of characteristic data, and they address the quality of property values that are exchanged within master data messages.

ISO 8000-110 includes the following:

— requirements for master data messages, which are used to exchange item characteristic data:

  — requirements adherence to a formal syntax;

  — requirements semantic encoding;

  NOTE Semantic encoding is the technique of replacing natural language terms in a message with identifiers that reference data dictionary entries.

  — requirements conformance to a customer data specification;

  — requirements business model.

ISO 8000-120 is an optional addition to Part 110 that includes the following:

— requirements for capture and exchange of data provenance information;

— data model for data provenance information.

ISO 8000-130 is an optional addition to Part 120 that includes the following:

— requirements for capture and exchange of data accuracy information in the form of statements and assertions of data accuracy;

— conceptual data model for data accuracy information in the form of statements and assertions of data accuracy.

ISO 8000-140 is an optional addition to Part 120 that includes the following:

— requirements for capture and exchange of data completeness information in the form of statements and assertions of data completeness;

— conceptual data model for data completeness information in the form of statements and assertions of data completeness.

ISO/TS 8000-150 includes the following:

— fundamental principles of master data quality management, and requirements for implementation, data exchange and provenance;

— an informative framework that identifies processes for data quality management.

Annex A
(normative)

Document identification

To provide for unambiguous identification of an information object in an open system, the object identifier

{ iso standard 8000 part (100) version (1) }

is assigned to this part of ISO 8000. The meaning of this value is defined in ISO/IEC 8824-1, and is described in ISO 10303-1.

Annex B
(informative)

Categories of items

B.1 General

Identifiers play a crucial role in supply chain management and product lifecycle support. There are three distinct categories of items, each with its own kind of identifiers:

— physical objects: asset tracking numbers and serial numbers;

— items of production: part or model numbers;

— items of supply: stock numbers.

B.2 Physical objects

B.2.1 Asset tracking or serial numbers

An asset tracking number or a serial number is a unique number given to a single physical object. If an item has a depreciable value or warranty, if it is taxed, or needs to be tracked, it will have an asset tracking number or serial number. While serial numbers are most common, they are not always assigned by the original manufacturer. There are several mandated and volunteer schemes to ensure that serial numbers remain unique within a domain or a period of time to ensure that no two items share the same number. If a serial number has not been assigned by a manufacturer, then a supplier or a buyer will affix an asset tracking number. Serial numbers are traditionally generated as sequential numeric or alphanumeric identifiers and the assigning organization will maintain a database that commonly tracks the origin and the current owner of the item.

Serial numbers are traditionally generated as sequential numeric or alphanumeric identifiers and the assigning organization will maintain a database that commonly tracks the origin and often the current owner of the item.

B.2.2 Batch numbers

A variation of a serial number is a batch number. Although a batch number is linked to a group of items or an amount of substance rather than to an instance of an item, batch numbers are used like a serial number for quality control and warranty purposes. Batch numbers are most often found on perishable items such as food or drugs or on consumable items for quality assurance and traceability purposes. The Item Unique Identification (IUID) is an example of a universal asset number

B.3 Items of production

B.3.1 Product, part and model numbers

A product number designates a type of item or substance that a manufacturer makes, has made, or plans to make. A part or model number is a kind of product number that a manufacturer or supplier uses to designate a group of discrete items that are considered interchangeable within a particular application. The main purpose of product, part or model numbers is to support sales and marketing, and they are the preferred numbers used in reordering. Manufacturers and suppliers often use different product, part or model numbers for identical substances or items designed to move through different channels of distribution. This is particularly common in the retail electronics industry (for example,

identical television models from the same manufacturer may have different model numbers depending on the retail outlets through which they are sold). It is also common in the oil industry, where the same underlying product is sold under different brands into different markets.

Part numbers are often designed to include some form of classification and often contain coded characteristics of the item. While part numbers are not necessarily unique, it is not unusual for companies to use part or model numbers as brands.

Several initiatives are designed to create universal part numbers. Most of these consist in adding a prefix that uniquely identifies the manufacturer or supplier who issued the number. The most common are bar codes such as the Universal Product Code (UPC) or its replacement, the Global Trade Identification Number (GTIN), issued by GS1, formerly a joint effort between the US Uniform Code Council (UCC) and European Article Number (EAN) International. The basic principal of the UCC/EAN UPC and GTIN numbers is central control of globally unique manufacturer or supplier prefixes associated with an understanding that the manufacturer or supplier controlled suffix should be unique to that manufacturer or supplier. There are other such initiatives including prefixing the part number with a commercial business identifier such as a Data Universal Numbering System (DUNS) number.

Although many items may share the same product or model number, each serial number will be unique to one item. There is a one-to-many relationship between product numbers and serial numbers (items with the same product number will have different serial numbers).

B.3.2 Properties of an item of production

The properties of an item of production consist of the following three types:

— reference numbers such as the part number, model number or procedure code. These references uniquely identify an item from the perspective of a manufacturer or supplier and are most commonly associated with a specification;

— dynamic properties such as price and availability (including location);

— properties of a subjective nature that are not suitable for use by a buyer for the purpose of objective differentiation of the fit, form or function of the item.

B.4 Items of supply

B.4.1 Stock numbers

The stock number, inventory number, or stockkeeping unit is a number issued by the buyer not only to track internal inventory but also as the primary method to support interoperability and competitive sourcing. As manufacturers and suppliers are also buyers, they also assign stock numbers to what they buy, assemble or make, and their in house stock numbers often become their outbound part numbers.

Most inventory management systems will use the stock number to link multiple sources of supply, either multiple suppliers for the same manufacturer part number or alternative manufacturers and part numbers. It is this concept of substitution that differentiates the stock number from the part number.

B.4.2 Properties of an item of supply

The properties of an item of supply represent the differentiating characteristics of an item. These consist of the properties that are descriptive of the fit, form, or function of an item and allow the item to be differentiated from others. Differentiating properties are commonly units of measures or enumerated values although it is possible for freeform text to constitute a differentiating property.

EXAMPLE The text in a road sign.

Bibliography

[1] ISO 3534-2:2006, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

[2] ISO/TS 8000-1, *Data quality — Part 1: Overview*

[3] ISO 8000-110, *Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification*

[4] ISO 8000-120, *Data quality — Part 120: Master data: Exchange of characteristic data: Provenance*

[5] ISO 8000-130, *Data quality — Part 130: Master data: Exchange of characteristic data: Accuracy*

[6] ISO 8000-140, *Data quality — Part 140: Master data: Exchange of characteristic data: Completeness*

[7] ISO/TS 8000-150, *Data quality — Part 150: Master data: Quality management framework*

[8] ISO 10303-1, *Industrial automation systems and integration — Product data representation and exchange — Part 1: Overview and fundamental principles*

[9] ISO/IEC 8824-1, *Information technology — Abstract Syntax Notation One (ASN.1) — Part 1: Specification of basic notation*

ICS 25.040.40

Price based on 14 pages