FINAL
DRAFT

INTERNATIONAL STANDARD

ISO/FDIS
8000-110

Data quality —

Part 110:
Master data — Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification

iTeh STANDARD PREVIEW

Qualité des données —
(standards.iteh.ai)

Partie 110: Données permanentes — Échange des données caractéristiques: Syntaxe, sémantique, encodage et conformité aux spécifications de données
ISO/FDIS 8000-110
https://standards.iteh.ai/catalog/standards/sist/03f1b42-949a-4bb0-88bc-cefad2bba1b8/iso-fdis-8000-110

Reference number
ISO/FDIS 8000-110:2021(E)

© ISO 2021

iTeh STANDARD PREVIEW
(standards.iteh.ai)

COPYRIGHT PROTECTED DOCUMENT

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 of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 184, *Automation systems and integration*, Subcommittee SC 4, *Industrial data*.

This second edition cancels and replaces the first edition (ISO 8000-110:2009), which has been technically revised.

The main changes compared to the previous edition are as follows:

— broken Uniform Resource Locators have been removed;

— normative references, figures and tables have been updated;

— the term "data value" has been replaced with the term "value tuple";

— the terms "property value" and "property value pair" have been replaced with the term "property-value tuple";

— an Annex B has been added which references an example of a schema that can be used to exchange master data that is characteristic data;

— editorial corrections to language, grammar and document layout were made.

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

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Digital data delivers value by enhancing all aspects of organizational performance including:

— operational effectiveness and efficiency;

— safety;

— reputation with customers and the wider public;

— compliance with statutory regulations;

— innovation;

— consumer costs, revenues and stock prices.

In addition, many organizations are now addressing these considerations with reference to the United Nations Sustainable Development Goals[1].

The influence on performance originates from data being the formalized representation of information. This information enables organizations to make reliable decisions. This decision making can be performed by human beings directly and also by automated data processing including artificial intelligence systems.

Through widespread adoption of digital computing and associated communication technologies, organizations become dependent on digital data. This dependency amplifies the negative consequences of lack of quality in this data. These consequences are the decrease of organizational performance.

The biggest impact of digital data comes from two key factors:

— the data having a structure that reflects the nature of the subject matter;

EXAMPLE 1   A research scientist writes a report using a software application for word processing. This report includes a table that uses a clear, logical layout to show results from an experiment. These results indicate how material properties vary with temperature. The report is read by a designer, who uses the results to create a product that works in a range of different operating temperatures.

— the data being computer processable (machine readable) rather than just being for a person to read and understand.

EXAMPLE 2   A research scientist uses a database system to store the results of experiments on a material. This system controls the format of different values in the data set. The system generates an output file of digital data. This file is processed by a software application for engineering analysis. The application determines the optimum geometry when using the material to make a product.

ISO 9000 explains that quality is not an abstract concept of absolute perfection. Quality is actually the conformance of characteristics to requirements. This actuality means that any item of data can be of high quality for one use but not for another. This difference occurs when the requirements are different between the two uses.

EXAMPLE 3   Time data are processed by calendar applications and also by control systems for propulsion units on spacecraft. This data includes start times for meetings in a calendar application and activation times in a control system. These start times require less precision than the activation times.

The nature of digital data is fundamental to establishing requirements that are relevant to the specific decisions made by an organization.

EXAMPLE 4   ISO/TS 8000-1 identifies that data has syntactic (format), semantic (meaning) and pragmatic (usefulness) characteristics.

---

1)   https://sdgs.un.org/goals

To support the delivery of high-quality data, the ISO 8000 series addresses:

— data governance, data quality management and maturity assessment;

EXAMPLE 5   ISO 8000-61 specifies a process reference model for data quality management.

— creating and applying requirements for data and information;

EXAMPLE 6   ISO 8000-110 (this document) specifies how to exchange characteristic data that is master data.

— monitoring and measuring data and information quality;

EXAMPLE 7   ISO 8000-8 specifies approaches to measuring data and information quality.

— improving data and, consequently, information quality;

EXAMPLE 8   ISO/TS 8000-81 specifies an approach to data profiling, which identifies opportunities to improve data quality.

— issues that are specific to the type of content in a data set.

EXAMPLE 9   ISO/TS 8000-311 specifies how to address quality considerations for product shape data.

Data quality management covers all aspects of data processing, including creating, collecting, storing, maintaining, transferring, exploiting and presenting data to deliver information.

Effective data quality management is systemic and systematic, requiring an understanding of the root causes of data quality issues. This understanding is the basis for not just correcting existing nonconformities but also implementing solutions that prevent future reoccurrence of those nonconformities.

EXAMPLE 10   If a data set includes dates in multiple formats including "yyyy-mm-dd", "mm-dd-yy" and "dd-mm-yy" then data cleansing can correct the consistency of the values. Such cleansing, however, requires additional information to resolve ambiguous entries (such as, "04-05-20"). The cleansing also cannot address any process issues and people issues, including training, that have caused the inconsistency.

As a contribution to this overall capability of the ISO 8000 series, this document supports the creation and exchange of high-quality data. This document contains requirements necessary but not sufficient to achieve data quality with respect to the exchange of master data. The requirements do not cover issues such as addressing the accuracy, provenance and completeness of master data. These issues need to be part of an overall data quality strategy adopted by each organization.

"Organization" does not necessarily mean a single, complete company or corporation. The organization can be a subdivision or branch that covers some distinct area of business operation.

When different business units of a company exchange master data or when a business unit exchanges master data with headquarters, these business units are organizations for the purposes of this document.

Organizations can use this document on its own or in conjunction with other parts of the ISO 8000 series.

This document supports activities that affect:

— one or more information systems;

— data flows within the organization and with external organizations;

— any phase of the data life cycle.

By implementing parts of the ISO 8000 series to improve organizational performance, an organization achieves the following benefits:

— objective validation of the foundations for digital transformation of the organization;

— a sustainable basis for data in digital form becoming a fundamental asset class the organization relies on to deliver value;

— securing evidence-based trust from other parties (including supply chain partners and regulators) about the repeatability and reliability of data and information processing in the organization;

— portability of data with resulting protection against loss of intellectual property and reusability across the organization and applications;

— effective and efficient interoperability between all parties in a supply chain to achieve traceability of data back to original sources;

— readiness to acquire or supply services where the other party expects to work with common understanding of explicit data requirements.

ISO/TS 8000-1 provides a detailed explanation of the structure and scope of the whole ISO 8000 series.

ISO 8000-2[2]) specifies the single, common vocabulary for the ISO 8000 series. This vocabulary is ideal reading material by which to understand the overall subject matter of data quality. ISO 8000-2 presents the vocabulary structured by a series of topic areas (for example, terms relating to quality and terms relating to data and information).

The following areas are outside the scope of this document and are covered in other international standards:

— ISO 8000-120 addresses the capture and exchange of data provenance information.

iTeh STANDARD PREVIEW
— ISO 8000-130 addresses the representation and exchange of information about the accuracy of master data that consists of characteristic data. (standards.iteh.ai)

— ISO 8000-140 addresses the representation and exchange of information about the completeness of master data that consists of characteristic data.

ISO/FDIS 8000-110
https://standards.iteh.ai/catalog/standards/sist/63f91b42-949a-4bb0-88bc-
— The ISO 9735 series, the ISO 13584 series, the ISO 15926 series and the ISO 22745 series specify formats that enable exchange of master data.

Annex A of this document contains an identifier that unambiguously identifies this document in an open information system.

Annex B of this document references an example of a schema to enable exchange of master data that is characteristic data.

---

2) The content is available on the ISO Online Browsing Platform. https://www.iso.org/obp

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Data quality —

Part 110:
Master data — Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification

1 Scope

This document specifies requirements for the exchange of messages that contain master data consisting of characteristic data. These requirements can be checked by computer. The messages are suitable for exchange between organizations and between systems.

EXAMPLE 1 A supplier sends a message to a customer. The message contains characteristic data describing an item that the customer is considering buying.

The following are within the scope of this document:

— conformance of master data messages to a formal syntax;

— semantic encoding of master data messages;

— conformance of master data messages to data specifications;

— requirements on access to the data dictionaries that enable decoding of master data messages.

The following are outside the scope of this document (see the Introduction for references to international standards that cover these areas):

— master data that is not characteristic data;

— data that is not in messages;

— messages that do not exchange master data between organizations or systems;

— recording the provenance of master data;

— accuracy of master data;

— exchange of data that is not master data;

— management of master data internally within an organization;

— data within an organization's enterprise resource planning or product data management system;

— making backup copies of data files containing master data;

— quality of data dictionaries;

— a specific formal syntax for the exchange of master data.

The requirements in this document are considered necessary but not sufficient to achieve data quality with respect to exchange of master data. Issues such as the accuracy and provenance of master data also need to be addressed as part of an overall data quality strategy.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

4 Fundamental concepts and assumptions

While the bit is the fundamental building block of electronically stored data, the property-value tuple is the fundamental building block of electronically stored characteristic data.

A value tuple by itself lacks the context to ensure unambiguous understanding of what the value represents. The context can be represented by a property label.

EXAMPLE 1   ISO 3601-1 specifies the inside diameters, cross-sections, tolerances and designation codes for O-rings used in fluid power systems for general industrial and aerospace applications. O-Rings are endless, ring-shaped sealing elements with a circular cross-section that are manufactured from elastomer materials or PTFE. The designation of an O-Ring is derived from the dimensions of its inner diameter and cord width in millimetres, along with the name of the material and its Shore A hardness (e.g, "100,00x2,65. NBR 70").

In this example, the full textual description for the O-Ring would be an O-Ring, part number 100,00x2,65 NBR 70 having an inside diameter of 100 millimetres, a cross-section of 2,65 millimetres, with a shore A hardness of 70 on the durometer scale, manufactured using nitrile-butadiene rubber (NBR), coloured black, that can be used in an operating temperature range of −30 °C, to +70 °C. The O-Ring sizes and tolerances conform to ISO 3601-1.

This natural language statement of the characteristic data can be decomposed into the elements shown in Table 1. The first column comprises the property labels; the second column comprises the value tuples.

Table 1 — Example characteristic data for an O-ring

Property label	Value tuple
INNER DIAMETER	100 MILLIMETRES
CROSS SECTION	2,65 MILLIMETRES
SHORE HARDNESS RATING	70 DUROMETER
MATERIAL	NITRILE-BUTADIENE RUBBER (NBR)
COLOUR	BLACK
OPERATING TEMPERATURE RANGE	−30 °C TO +100 °C
MANUFACTURER CODE	100,00X2,65 NBR 70
STANDARD	ISO 3601-1

Using property labels has the following disadvantages:

— the labels are in a single language, which can prevent understanding of the data by speakers of other languages;

— the labels are susceptible to variations in spelling, letter case, punctuation, etc., making it difficult for a computer application to understand the characteristic data;

— the labels are not defined, which can result in differing interpretations between the sender and receiver.

To address these problems, this document requires the use of property-value tuples when exchanging master data messages that contain characteristic data. Each property-value tuple consists of a value tuple and, to establish the context of that value, a reference to a data dictionary entry that specifies the property. The entry contains terms and definitions of the property in one or more languages.

It follows that, as property-value tuples are the fundamental building blocks of electronically stored information, the quality of the property-value tuples is one of the determinants of the quality of the information.

Semantic encoding produces messages that contain property-value tuples where each reference to a data dictionary entry is in the form of an unambiguous identifier.

Use of a data dictionary is the foundation for semantic encoding, which, in combination with identifying the applicable syntax for a data set, establishes that data as being portable. Portable data can pass from one system to another reliably because the receiver can guarantee being able to process the format and the meaning of the data.

To ensure data is portable without subsequent claims of ownership by third parties, organizations choose a data dictionary that has appropriate licence conditions.

The remainder of this document specifies how to create portable data that conforms to applicable data specifications.

For the parts of the ISO 8000 series where those parts address master data, a foundational data architecture supports the exchange of data without loss of content or meaning (see Figure 1).
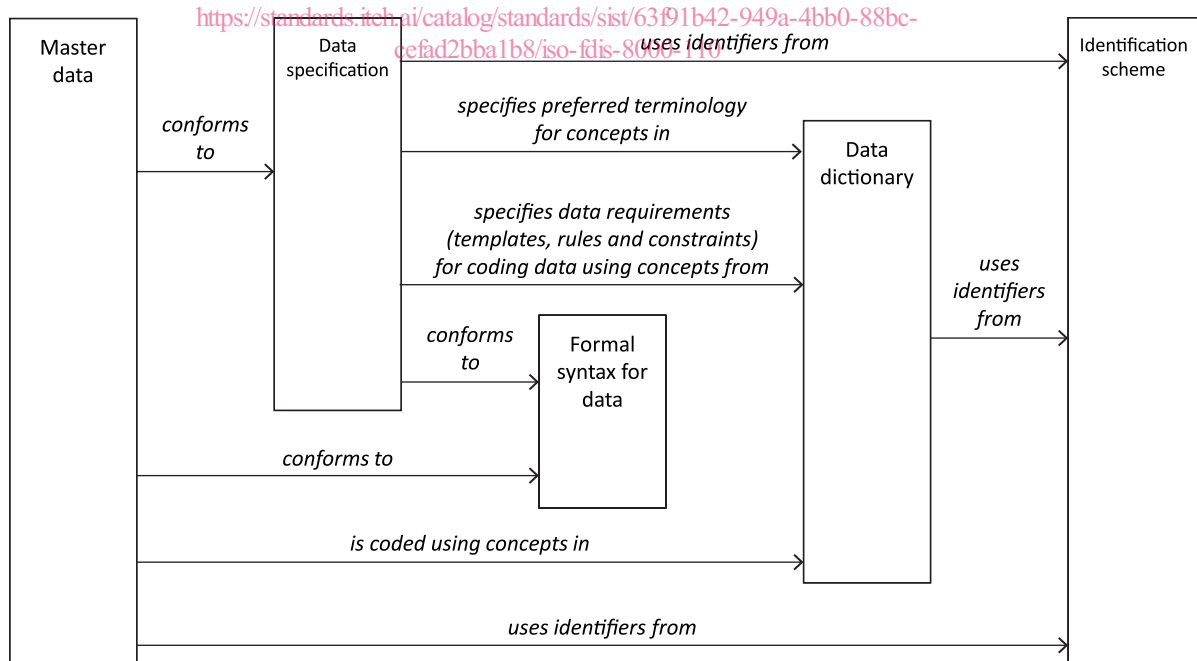

Figure 1 — Data architecture for master data as specified by the ISO 8000 series