भारतीय मानक **IS/ISO 8000-1 : 2023**

Indian Standard

आंकड़ो की गुणवत्ता

भाग 1 अवलोकन

Data Quality

Part 1 Overview

ICS 25.040.40

भारतीय मानक ब्यूरो
BUREAU OF INDIAN STANDARDS
मानक भवन, 9 बहादुर शाह ज़फर मार्ग, नई दिल्ली - 110002
MANAK BHAVAN, 9 BAHADUR SHAH ZAFAR MARG
NEW DELHI - 110002
www.bis.gov.in www.standardsbis.in

November 2023 **Price Group 10**

NATIONAL FOREWORD

This Indian Standard which is identical to ISO 8000-1 : 2022 'Data quality — Part 1: Overview' issued by the International Organization for Standardization (ISO) was adopted by the Bureau of Indian Standards on recommendation of the Industrial Automation Systems and Robotics Sectional Committee had been approved by the Production and General Engineering Division Council.

Other parts in this series are:

Part 2	Vocabulary
Part 8	Information and data quality — Concepts and measuring
Part 60	Data quality management — Overview
Part 61	Data quality management — Process reference model
Part 62	Data quality management — Organizational process maturity assessment: Application of standards relating to process assessment
Part 63	Data quality management — Process measurement
Part 64	Data quality management — Organizational process maturity assessment: Application of the test process improvement method
Part 65	Data quality management — Process measurement questionnaire
Part 66	Data quality management — Assessment indicators for data processing in manufacturing operations
Part 81	Data quality assessment — Profiling
Part 82	Data quality assessment — Creating data rules
Part 100	Master data — Exchange of characteristic data: Overview
Part 110	Master data — Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification
Part 115	Master data: Exchange of quality identifiers — Syntactic, semantic and resolution requirements
Part 116	Master data — Exchange of quality identifiers — Application of ISO 8000-115 to authoritative legal entity identifiers
Part 120	Master data — Exchange of characteristic data — Provenance
Part 130	Master data — Exchange of characteristic data — Accuracy
Part 140	Master data — Exchange of characteristic data — Completeness
Part 150	Data quality management — Roles and responsibilities
Part 311	Guidance for the application of product data quality for shape (PDQ-S)

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

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 these data. These consequences are the decrease of organizational performance.

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

a) the data having a structure that reflects the nature of the subject matter; and
b) the data being computer processable (machine readable) rather than just being for a person to read and understand.

IS/ISO 9000 explains that quality is not an abstract concept of absolute perfection. Quality is the conformance of characteristics to requirements. This actuality means that any item of data can be of high quality for one purpose but not for a different purpose. The quality is different because the requirements are different between the two purposes.

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.

(*Continued on third cover*)

Contents

Page

Introduction

Digital data deliver 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[2)]. This information enables organizations to make reliable decisions. Such 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 these 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 purpose but not for a different purpose. The quality is different because the requirements are different between the two purposes.

EXAMPLE 3 Time data are processed by calendar applications and also by control systems for propulsion units on spacecraft. These data include 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 8000-8 identifies that data have syntactic (format), semantic (meaning) and pragmatic (usefulness) characteristics.

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

2) ISO 8000-2 defines information as “knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning”.

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 specifies how to exchange characteristic data that are master data.

— monitoring and measuring information and data quality;

EXAMPLE 7 ISO 8000-8 specifies approaches to measuring information and data 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 requires additional information, however, 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 provides a detailed explanation of the structure and scope of the whole ISO 8000 series.

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 8000-2[3)] specifies the single, common vocabulary for the ISO 8000 series. This vocabulary is a foundation for understanding 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).

ISO has identified this document, ISO 8000-2 and ISO 8000-8 as horizontal deliverables[4)].

Annex A contains an identifier that conforms to ISO/IEC 8824-1. The identifier unambiguously identifies this document in an open information system.

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

4) Deliverable dealing with a subject relevant to a number of committees or sectors or of crucial importance to ensure coherence across standardization deliverables.

Indian Standard

DATA QUALITY

PART 1 OVERVIEW

1 Scope

This document provides an overview of the ISO 8000 series.

The following are within the scope of this document:

— stating the scope of the ISO 8000 series as a whole;

— establishing the principles of information and data quality;

— describing the path to data quality;

— describing the structure of the ISO 8000 series;

— providing a summary of the content of each part in the ISO 8000 series;

— establishing the relationship of the ISO 8000 series to other international standards.

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 terminology 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 Scope of the ISO 8000 series

The ISO 8000 series provides frameworks for improving data quality for specific kinds of data. The series defines which characteristics of data are relevant to data quality, specifies requirements applicable to those characteristics, and provides guidelines for improving data quality. The series is applicable within all stages of the data life cycle.

NOTE The ISO 8000 series can be used either in conjunction with or independently of standards for quality management systems.

The following are within the scope of the ISO 8000 series:

— general aspects of data quality, including principles, vocabulary and measurement of information and data quality;

— data governance;

— data quality management, including processes, roles, responsibilities and maturity assessment;

— data quality assessment, including profiling and data rules;

— quality of master data, including exchange of characteristic data and identifiers;

— quality of industrial data, including product shape data.

The following are outside the scope of the ISO 8000 series:

— quality of the things represented by data;

EXAMPLE 1 AF Industries makes fasteners. AF publishes an electronic catalogue of its products. The quality of the catalogue (the data) is within the scope of the ISO 8000 series. The quality of the fasteners (the things represented by the data) is outside of the scope of the ISO 8000 series.

— quality management principles;

EXAMPLE 2 ISO 9000 identifies eight quality management principles: customer focus, leadership, engagement of people, process approach, improvement, evidence-based decision making and relationship management.

— systems and software product quality.

EXAMPLE 3 ISO/IEC 25000, ISO/IEC 25010, ISO/IEC 25012 and ISO/IEC 25024 address systems and software product quality requirements and evaluation.

The ISO 8000 series contains requirements that are intended to be applicable to:

— all organizations, regardless of type and size;

— organizations at each point in the data supply chain.

5 Principles of information and data quality

The following principles of information and data quality underlie the ISO 8000 series:

— data are the reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing;

— an agreed level of data quality enables the right people to make the right decisions at the right time;

NOTE 1 Making the right decisions validates that the data are fit for purpose. Decision making can depend on many different characteristics including location and sequence.

NOTE 2 Agreement is necessary between the stakeholders participating in and affected by the decision.

— effective data quality management builds on the fundamental concepts and principles of ISO 9000;

— data quality is a function of the inherent characteristics of the data under consideration;

EXAMPLE 1 A data set contains only characters from the character set specified by ISO/IEC 10646. This fact is an inherent characteristic of the data.

EXAMPLE 2 One item in a data set has a length of 13 (i.e. is represented by 13 characters). This length is an inherent characteristic of the data.

EXAMPLE 3 One column in a database table contains only values that are members of the set {"cm", "inches", "mm", "m"}. This fact is an inherent characteristic of the data.

EXAMPLE 4 ISO/IEC 27001 addresses confidentiality, integrity and availability of data as the primary considerations for management of information security. Confidentially, integrity and availability are functions of the whole system (including hardware, software and people) and, thus, are not inherent characteristics of the data.

— organizations gain greater value when digital data are computer processable;

— sustained co-ordinated activity by an organization protects and realises the value of data as an asset.

NOTE 3 This activity requires levels of investment that are appropriate to the value delivered by the data. The co-ordination addresses the influence that everyone in the organization has on the quality and value of data.

6 The ISO 8000 series path to data quality

The ISO 8000 series specifies a comprehensive capability to achieve and sustain high quality data. This capability (see Figure 1):

— builds on the principles of data quality (see Clause 5);

— responds to the information need arising from the decision-making requirements of an organization;

— consists of processes that deliver effective data quality management and achieve governance and assurance;

NOTE 1 These processes follow the fundamental structure of the Plan-Do-Check-Act cycle as also adopted by ISO 9001.

— makes use of data specifications to address information need and to establish a rigorous basis for determining that data meet applicable requirements in respect of syntactic, semantic and pragmatic considerations;

— supports verification and validation of conformance to requirements;

— ensures alignment with the strategic objectives of the organization;

— delivers sustainable organizational and behavioural changes to improve data quality;

— is applicable to any type of data.

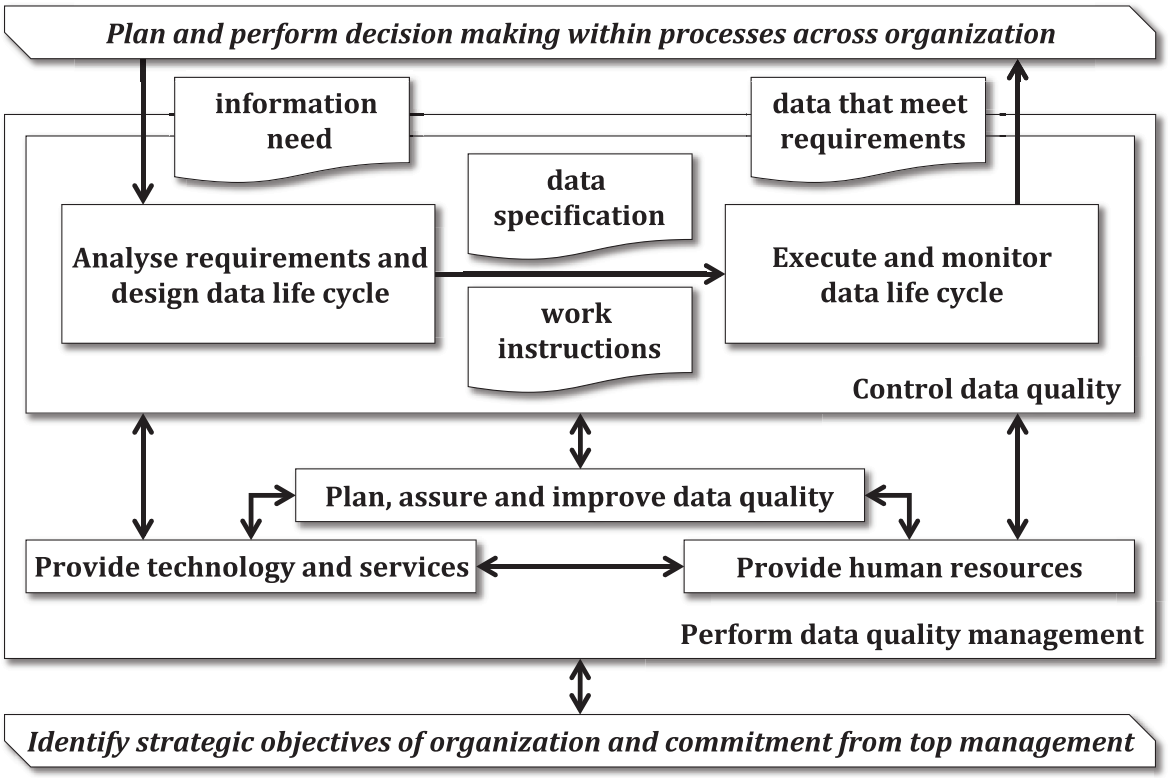

Figure 1 — ISO 8000 series capability to achieve and sustain high quality data

Although the ISO 8000 series is applicable to any type of data, the series does address some of the different considerations that distinguish between types. These considerations include the:

— role of the data;

EXAMPLE 1 Roles include, but are not limited to: master data; transaction data; reference data.

EXAMPLE 2 The role of master data is to describe the entities that are both independent and fundamental for an organization. The data are referenced in order to perform transactions. The entities include customers, products, employees, materials, suppliers, services, shareholders, facilities, equipment, rules and regulations.

— component within the data architecture;

EXAMPLE 3 Components include, but are not limited to: dictionary; exchange file; database, data schema.

EXAMPLE 4 A dictionary consists of data that uniquely identify and describe entities such as part classifications, types of property and units of measurement. This dictionary is a component of the data architecture within an organization. This component enables consistent understanding of other data managed by the organization.

— scope of information represented by the data;

EXAMPLE 5 Scope includes, but is not limited to: products (for example, product shape data); projects; financial records; identifiers; performance characteristics; physical characteristics; safety. These scopes are not necessarily distinct from each other.

EXAMPLE 6 The subject matter of product data is all the different aspects that determine the functions, performance and how to conduct activities on the product (including creation, maintenance and disposal).

— intended area of application of the data;

EXAMPLE 7 Areas of application include, but are not limited to: industry; healthcare; banking.

EXAMPLE 8 Industrial data enables an organization to develop, manufacture and distribute a range of products to fulfil customer demand.

— structure of the data.

NOTE 2 The terms “structured data” and “unstructured data” are sometimes used to suggest a clear, binary distinction in the structure of data. This distinction is, however, a simplification of typical digital data sets, which actually combine structured and unstructured elements.

EXAMPLE 9 In an equipment database with less structure, an organization describes each item of equipment with a single free-text entry that covers all the characteristics of the item and is unable to enforce any standardisation of how different persons choose to describe the item. In an equipment database with more structure, the organization is able to record discrete values for individual characteristics of each item. These characteristics could include length, width and height, each with an explicit unit of measurement to ensure consistency and interpretability of the database entry for each item.

In the above list of considerations, each item is independent of the other items.

EXAMPLE 10 An item of master data (role) that is highly structured product data (scope) and forms part of a wider set of industrial data (intended application).

7 Structure of the ISO 8000 series

7.1 Overview of the structure

The ISO 8000 series consists of groups of standards that address the following:

— general aspects of data quality (see 7.2);

— data governance (see 7.3);

— data quality management (see 7.4);

— data quality assessment (see 7.5);

— quality of master data (see 7.6);

— quality of industrial data (see 7.7).

Each group consists of one or more parts, which address specific aspects of data quality.

Table 1 lists the current parts of the ISO 8000 series.

Organizations can:

— use individual parts of the ISO 8000 series to address specific issues;

— implement multiple parts to achieve a more comprehensive capability for data quality.

Table 1 — Parts of the ISO 8000 series

Group of parts	Part	Description
General aspects of information and data quality	ISO 8000-1	Overview of information and data quality
	ISO 8000-2	Vocabulary for information and data quality
	ISO 8000-8	Measuring and reporting information and data quality
Data governance	ISO 8000-51 [a]	Exchange of data policy statements
Data quality management	ISO/TS 8000-60	Overview of data quality management
	ISO 8000-61	Process reference model for data quality management
	ISO 8000-62	Determining maturity of data quality management by applying standards relating to process assessment
	ISO 8000-63	Process measurement approach
	ISO 8000-64	Determining maturity of data quality management by applying the Test Process Improvement Method
	ISO/TS 8000-65	Process measurement questionnaire for data quality management
	ISO 8000-66	Maturity of data quality management for manufacturing operations
	ISO 8000-150	Roles and responsibilities for data quality management
Data quality assessment	ISO/TS 8000-81	Data profiling
	ISO/TS 8000-82 [b]	Creating data rules
Quality of master data	ISO 8000-100	Overview of master data quality
	ISO 8000-110	Requirements for the exchange of characteristic data
	ISO 8000-115	Exchange of identifiers
	ISO 8000-116	Authoritative legal entity identifiers
	ISO 8000-120	Stating provenance when exchanging characteristic data
	ISO 8000-130	Stating accuracy when exchanging characteristic data
	ISO 8000-140	Stating completeness when exchanging characteristic data
Quality of industrial data	ISO/TS 8000-311	Product data quality for shape

[a] Under preparation. Stage at the time of publication: ISO/DIS 8000-51.

[b] Under preparation. Stage at the time of publication: ISO/PRF TS 8000-82.

7.2 General aspects of information and data quality

7.2.1 Group purpose and constituents

The purpose of the general aspects group of parts is to enable organizations to understand the purpose and scope of the ISO 8000 series. This understanding ensures each organization is able to choose which parts of the ISO 8000 series are relevant to the circumstances of the organization. This choice supports the implementation of sustainable information and data quality.

The purpose of the general aspects group of parts also includes enabling organizations to implement effective identification of the characteristics that determine the quality of information and data. This effectiveness ensures organizations achieve coherent and repeatable quantification of quality issues in data sets.

The following topics are addressed by the parts in the general aspects group:

— an overview of information and data quality and how the ISO 8000 series deals with the topic (the contents of this document);

— a vocabulary for information and data quality (see 7.2.2 and ISO 8000-2);

— measuring and reporting information and data quality (see 7.2.3 and ISO 8000-8).

7.2.2 ISO 8000-2

ISO 8000-2 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.

EXAMPLE Two of the subclauses in ISO 8000-2 are “Terms relating to quality” and “Terms relating to data and information”.

7.2.3 ISO 8000-8

ISO 8000-8 specifies prerequisites for measuring information and data quality when executed within processes and systems for quality management. These prerequisites include:

— understanding the nature of information and data quality, including the existence of syntactic, semantic and pragmatic quality;

— a structured way to plan and perform information and data quality measurements;

— an approach to reporting information and data quality measurements.

ISO 8000-8 is applicable to all types of:

— organization;

— information and data;

— data processing;

— technology used for that processing.

7.3 Data governance

7.3.1 Group purpose and constituents

The purpose of the data governance group of parts is to enable an organization to implement effective development and enforcement of policies related to the management of data. This effectiveness ensures organizations perform data processing that:

— aligns with strategic objectives of the organization;

— satisfies the requirements of all stakeholders, including those external to the organization.

EXAMPLE The General Data Protection Regulation [54] specifies data protection principles, rights and obligations. This regulation applies to most organizations operating within the territories of the European Union.

The following topics are addressed by the parts in the data governance group:

— exchange of data policy statements (see 7.3.2 and ISO 8000-51).

7.3.2 ISO 8000-51

ISO 8000-51 specifies requirements that support the exchange of data governance policy statements and the automation of testing the compliance of data sets to applicable policy statements. These requirements cover the syntax and semantics of identifiers for organizations issuing data governance policy statements and for those statements.

7.4 Data quality management

7.4.1 Group purpose and constituents

The purpose of the data quality management group of parts is to enable organizations to implement effective planning, control, assurance and improvement of the quality of data. This effectiveness ensures organizations perform data processing that:

— consistently generates or sustains data that meet requirements;

— delivers value to one or more classes of stakeholder.

NOTE Some stakeholders will be direct users of the data. Other stakeholders will benefit because someone else has exploited the data.

EXAMPLE A government analyses data to improve delivery of social services to citizens.

The following topics are addressed by the parts in the data quality management group:

— an overview of data quality management (see 7.4.2 and ISO/TS 8000-60);

— a process reference model for data quality management (see 7.4.3 and ISO 8000-61);

— applying standards relating to process assessment to perform the process maturity assessment of data quality management in an organization (see 7.4.4 and ISO 8000-62);

— a process measurement approach to data quality management (see 7.4.5 and ISO 8000-63);

— application of the Test Process Improvement method (see 7.4.6 and ISO 8000-64);

— a process measurement questionnaire for data quality management (see 7.4.7 and ISO/TS 8000-65);

— maturity of data quality management for manufacturing operations (see 7.4.8 and ISO 8000-66);

— roles and responsibilities for data quality management (see 7.4.9 and ISO 8000-150).

7.4.2 ISO/TS 8000-60

ISO/TS 8000-60 is an introduction to the group of standards that specify requirements applicable to data quality management. The introduction describes the core concepts applicable to data quality management and gives an overview of each part in the group. Those parts enable the implementation, assessment and improvement of data quality management.

7.4.3 ISO 8000-61

ISO 8000-61 specifies a process reference model for data quality management. This model supports assessing and improving the capability of the processes and increasing organizational maturity with respect to data quality management. The model describes each process in terms of a purpose, outcomes and activities. ISO 8000-61 also lists the fundamental principles of data quality management.

7.4.4 ISO 8000-62

ISO 8000-62 specifies how organizations can use a maturity model in assessing their process maturity with respect to data quality management as specified by ISO 8000-61 (see 7.4.3). This assessment requires the use of assessment indicators and can use the measurement stack specified by ISO 8000-63 (see 7.4.5) to determine these indicators. The maturity model conforms to ISO/IEC 33004.

7.4.5 ISO 8000-63

ISO 8000-63 specifies a process measurement approach that is appropriate for use when assessing process maturity. This approach can serve when an organization is looking to improve the maturity of

data quality management. The approach makes use of a structure for process measurement stacks that organizations can instantiate to measure the characteristics of processes for data quality management. This structure consists of goal, sub goal, question, indicator and metric. The instantiated stack consists of content that is determined by a chosen model for assessing the maturity of the processes under consideration.

7.4.6 ISO 8000-64

ISO 8000-64 specifies a procedure by which an organization can assess process maturity according to the specific priorities of the organization. This procedure provides a capability to assess and improve data quality management as specified by ISO 8000-61 (see 7.4.3). The procedure makes use of the Test Process Improvement method [55] and the measurement stack specified by ISO 8000-63 (see 7.4.5).

7.4.7 ISO/TS 8000-65

ISO/TS 8000-65 establishes a simple measurement method to evaluate the implementation of data quality management implementation by organizations. This method uses the processes from the reference model specified by ISO 8000-61 (see 7.4.3). The method poses a series of questions, each of which addresses one of the outcomes of the corresponding process from ISO 8000-61. The questions are applicable to all types of business process, technology, information system, data and data processing.

NOTE The term "business process" refers to the wider process within which the specific data processing is taking place. This wider process involves decision making that depends on the data. The process typically generates a product or service, although this can be for the internal purposes of an organization. This is also not necessarily for commercial gain in the case of government or other types of organization.

7.4.8 ISO 8000-66

ISO 8000-66 supports the application of ISO 8000-62 (see 7.4.4) to determine the process maturity of data quality management in manufacturing organizations. This support is provided by specifying assessment indicators for data processing in manufacturing operations management that is specified by IEC 62264-1. These indicators conform to ISO/IEC 33004.

7.4.9 ISO 8000-150

ISO 8000-150 addresses key considerations when establishing the roles and responsibilities necessary to deliver effective and efficient data quality management. These considerations are supported by a framework that links role levels to structured groups of responsibility and a model of operations to deliver data quality management. ISO 8000-150 also provides example scenarios for deployment of the framework. The role levels and responsibility groups are appropriate for all types of data and all types of organization.

7.5 Data quality assessment

7.5.1 Group purpose and constituents

The purpose of the data quality assessment group of parts is to enable organizations to implement effective identification and resolution of quality issues in data sets. This effectiveness ensures organizations:

— uncover previously hidden quality issues;

— prevent future recurrence of the same quality issues.

The following topics are addressed by the parts in the data quality assessment group of standards:

— an approach to data profiling (see 7.5.2 and ISO/TS 8000-81);

— creating data rules (see 7.5.3 and ISO/TS 8000-82).

7.5.2 ISO/TS 8000-81

ISO/TS 8000-81 specifies an approach to data profiling, which involves applying analysis techniques to data in actual use. This analysis generates a profile consisting of the structure, columns and relationships of the data. The profile provides the basis for identifying opportunities to improve data quality by establishing new explicit rules for the data. The approach also typically produces greater effect from repeated application to uncover issues progressively.

7.5.3 ISO/TS 8000-82

ISO/TS 8000-82 specifies how different data rules apply to various types of data. Such rules exist to sustain the integrity and reliability of data by capturing requirements into a form that can be processed by databases and other information systems. Each rule is able to support data quality assessment.

7.6 Quality of master data

7.6.1 Group purpose and constituents

The purpose of the master data group of parts is to enable organizations to implement effective processing of master data. This effectiveness ensures organizations:

— exploit or supply master data that meet a relevant range of applicable requirements;

EXAMPLE 1 Master data identifies and describes individuals, organizations, locations, goods, services, processes, rules and regulations.

— exchange data without loss of content or meaning.

In the master data group of parts, the foundation is a data architecture that supports such exchange (see Figure 2).

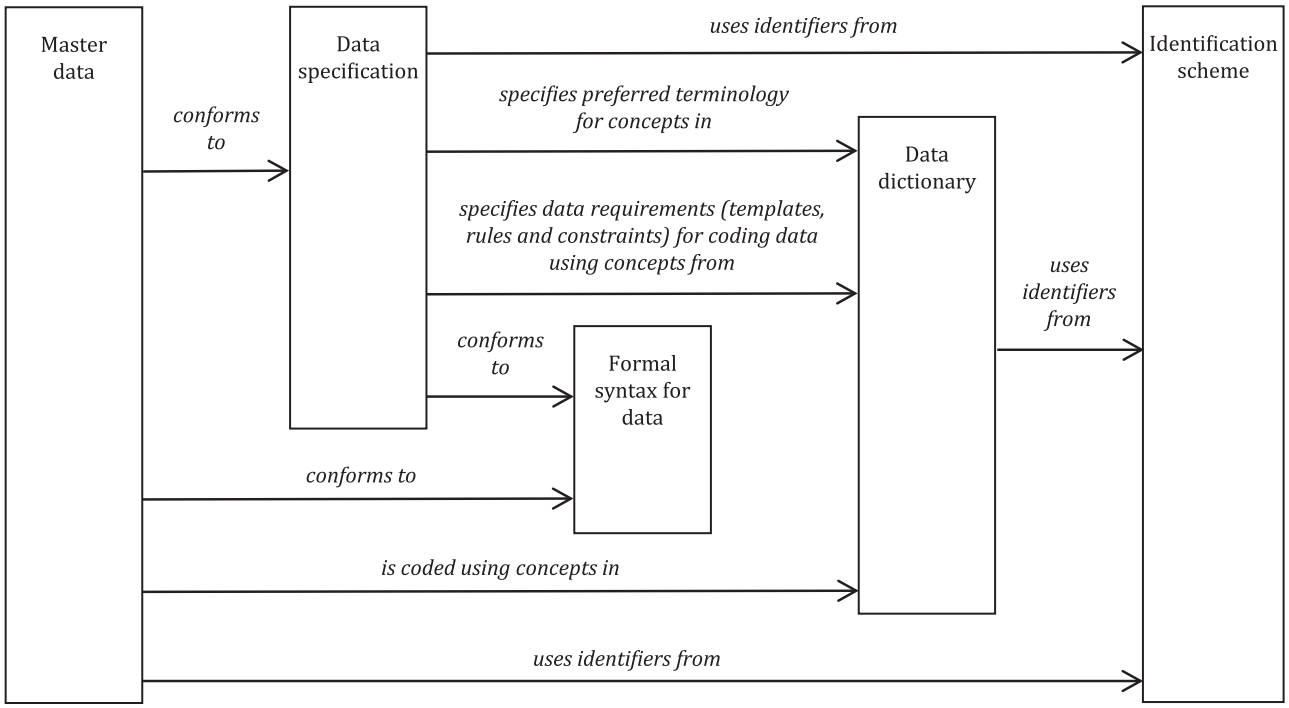

Figure 2 — Data architecture for master data as specified by the ISO 8000 series

The data architecture establishes the basis for consistent, high quality master data. This basis includes the data dictionary and data specifications as key elements

EXAMPLE 2 A data dictionary can take many forms including: a concept dictionary (see ISO/TS 29002-6[5]); an open technical dictionary (see ISO/TS 22745-10), a reference data library (see ISO/TS 15926-4) and a parts library (see the ISO 13584 series). In the case of ISO/TS 15926-4 and the ISO 13584 series, the data dictionary includes elements of the data specification.

EXAMPLE 3 A data specification can take many forms including: an identification guide (see ISO/TS 22745-30); a reference data item (see ISO/TS 15926-4); a product characterization class (see ISO 13584-42) and a class (see IEC 61360-4).

In addition to the data dictionary and data specifications, exchange of master data requires a formal syntax and an identification scheme.

EXAMPLE 4 Formal syntaxes include: eXtensible Markup Language (XML) (see ISO/TS 29002-10); the clear text encoding specified by the ISO 10303 series (see ISO 10303-21).

EXAMPLE 5 ISO/TS 29002-5 specifies an identification scheme in the form of an internationally recognized data identifier (IRDI). This identifier is a unique identifier that identifies administered items, each of which is either a concept or a concept information element. ISO/TS 29002-5 is specialized by ISO 22745-13, which uses the term “metadata object” to refer to administered items that are relevant to open technical dictionaries. ISO/TS 29002-5 enables interoperability between implementations of the ISO 13584 series for parts libraries and the ISO 22745 series for open technical dictionaries.

The following topics are addressed by the parts in the master data group of standards:

— an overview of master data quality and a data architecture for master data (see 7.6.2 and ISO 8000-100);

— requirements for the exchange of characteristic data (see 7.6.3 and ISO 8000-110);

— syntactic, semantic and resolution requirements for quality identifiers (see 7.6.4 and ISO 8000-115);

— an application of ISO 8000-115 to authoritative legal entity identifiers (see 7.6.5 and ISO 8000-116);

— provenance requirements for exchanging characteristic data (see 7.6.6 and ISO 8000-120);

— accuracy requirements for exchanging characteristic data (see 7.6.7 and ISO 8000-130);

— completeness requirements for exchanging characteristic data (see 7.6.8 and ISO 8000-140).

7.6.2 ISO 8000-100

ISO 8000-100 is an overview of the group of parts that address the quality of master data. This overview consists of an introduction to master data, a data architecture for master data, a high-level data model and a description of each of the parts in the group.

7.6.3 ISO 8000-110

ISO 8000-110 specifies requirements that can be checked by computer for the exchange of master data consisting of characteristic data. This exchange can take place between organizations and between software applications. These requirements enable organizations to manage the quality of master data. The requirements are independent of formal syntax and can be used by organizations to support implementation of any other standard that specifies a formal syntax for data sets.

The requirements in ISO 8000-110 apply to the exchange of master data messages that conform to a foundational data architecture for master data (see 7.6.1).

5) The ISO/TS 29002 series will be replaced by ISO 29002 (currently at DIS stage).

7.6.4 ISO 8000-115

ISO 8000-115 specifies requirements for identifiers that form part of an exchange of master data. These requirements supplement those of ISO 8000-110 (see 7.6.3).

Most commonly, an identifier is a reference to a data set managed by the owner of the identifier. This makes the identifier an alias for a master data record. Identifiers are widely exchanged by governments and commercial companies to refer to data that describe individuals, organizations, locations, goods, services, assets, processes, procedures, laws, rules and regulations.

EXAMPLE Types of identifier include vehicle registration number (license plate), vehicle identification number (VIN), driver's permit number, social security number, national identity card number, student number, employee number, passport number, tax identification number, Internet Protocol (IP) address, telephone number, e-mail address, domain name, part number, batch number, serial number, customer number, supplier number and concept identifier.

7.6.5 ISO 8000-116

ISO 8000-116 specifies the requirements for representing authoritative legal entity identifiers (ALEI). These requirements extend those of ISO 8000-115 (see 7.6.4).

Organizations often use proxy identifiers for individuals and organizations. These proxy identifiers are not issued by an organization that is the administrative agency of the government granting legal status to a physical or juridical person.

ISO 8000-116 enables organizations to formulate appropriate prefixes that conform to ISO 8000-115. These prefixes are the basis for open, globally unambiguous representation of authoritative legal entity identifiers.

The representation specified by ISO 8000-116 supports the supply chain reliably to obtain the legal name, date of formation and status of trading partners. This reliability arises from having explicit traceability to the administrative agency that granted each organization legal status under the authority of a government.

7.6.6 ISO 8000-120

ISO 8000-120 specifies requirements on the representation and exchange of information about the provenance of master data that consist of characteristic data. This information is the foundation for supporting claims of data accuracy (using ISO 8000-130, see 7.6.7) and data completeness (using ISO 8000-140, see 7.6.8).

ISO 8000-120 extends the requirements of ISO 8000-110 (see 7.6.3) but can also support the exploitation of any other standard that specifies a formal syntax for data sets.

7.6.7 ISO 8000-130

ISO 8000-130 specifies requirements on the representation and exchange of information about the accuracy of master data that consist of characteristic data. These requirements cover statements and assertions of this accuracy.

ISO 8000-130 extends the requirements of ISO 8000-120 (see 7.6.6).

7.6.8 ISO 8000-140

ISO 8000-140 specifies requirements on the representation and exchange of information about the completeness of master data that consist of characteristic data. These requirements cover statements and assertions of this completeness.

ISO 8000-140 extends the requirements of ISO 8000-120 (see 7.6.6).

7.7 Quality of industrial data

7.7.1 Group purpose and constituents

The purpose of the industrial data group of parts is to enable organizations to implement effective processing of data that represent various types of industrial information. This effectiveness ensures organizations:

— exploit or supply data that meet a relevant range of applicable requirements;

— deliver value through data-based decision making in industrial processes.

EXAMPLE Collaborative product development can involve significant rework of data if the creator of those data does not understand and address agreed requirements. Such product development makes use of a wide range of data supporting computer-aided design, computer-aided manufacturing, computer-aided engineering and product data management.

The following topics are addressed by the parts in the industrial data group of standards:

— use of product data quality for shape (see 7.7.2 and ISO/TS 8000-311).

7.7.2 ISO/TS 8000-311

ISO/TS 8000-311 provides guidance on how to use ISO 10303-59 to address the challenges of product data quality for shape (PDQ-S). This guidance explains the relationship of ISO 10303-59 to other standards that support the exchange, sharing and archiving of product data. The guidance explains how different circumstances can benefit from effective application of PDQ-S.

8 Relationship of the ISO 8000 series to other international standards

8.1 Overview of the relationship to other international standards

The ISO 8000 series:

— builds on existing approaches to quality management;

— does not replace detailed data requirements that exist in other standards;

— specifies requirements that complement any formal, systematic approach to activities involving data exploitation.

On this basis, the ISO 8000 series has implications for standards applicable to:

— quality management systems (see 8.2);

— other management systems (see 8.3);

— software quality (see 8.4);

— industrial data (see 8.5);

— other types of data (see 8.6).

8.2 Standards for quality management systems

The ISO 8000 series specifies an approach that addresses data quality in line with the requirements of ISO 9000, ISO 9001 and ISO 9004. These requirements are more generic than those in the ISO 8000 series.

The ISO 8000 series adopts the key vocabulary of ISO 9000. This vocabulary is the core for ISO 8000-2, which specifies a complete vocabulary for data quality (see 7.2.2).

ISO 9001 specifies requirements for a quality management system where an organization:

— needs to demonstrate a consistent ability to provide product that meets customer and applicable regulatory requirements;

— aims to enhance customer satisfaction through the effective application of the system.

ISO 9001 promotes the adoption of a process approach when developing, implementing and improving the effectiveness of a quality management system. This approach enhances customer satisfaction by meeting customer and applicable regulatory requirements.

Data are an output of all types of process in an organization. This output is fundamentally the same as any product and, therefore, subject to the general principles of quality management that are specified by ISO 9001. Data are, however, intangible, which does lead to some additional quality management considerations.

EXAMPLE Processes include those associated with business and manufacturing.

On this basis, the ISO 8000 series does not establish a new management system. The series, instead, extends and clarifies ISO 9001 for the case where data are the product.

ISO 9004 gives guidance on a wider range of objectives of a general quality management system than does ISO 8000. This range includes, in particular, the continual improvement of the overall performance, efficiency and effectiveness of organizations. ISO 9004 is a guide for organizations whose top management wishes to build on the requirements of the ISO 8000 series and pursue continual improvement of information performance. ISO 9004 is not, however, appropriate for certification or for contractual purposes.

ISO 9000, ISO 9001 and ISO 9004 focus exclusively on the process approach. This focus is narrower than the scope of the ISO 8000 series, which:

— promotes application of the process approach to data quality;

— identifies specific characteristics of data and explains how those characteristics affect quality;

— specifies requirements for exchanging data and the associated quality characteristics.

8.3 Standards for management systems other than quality management systems

Various standards address management systems other than quality management systems. For these standards, the fundamental connection to the ISO 8000 series is the role of data in decisions. In every management system, decisions are the means by which to ensure consistent and effective performance of an organization. The ISO 8000 series, thus, complements standards for management systems not addressing quality. The series does not, however, contain any requirements specific to any such standard.

EXAMPLE 1 ISO standards specify management systems that address topics including environmental performance, risk, information security and occupational health and safety management.

EXAMPLE 2 ISO 55001:2014, 7.5 specifies the information requirements that support assets, asset management, asset management systems and achieving organizational objectives.

8.4 Standards for software quality

The ISO 8000 series provides a focus on digital data. Such data only become accessible to users or other information technology devices by being part of a wider system. These systems contain software as a necessary component. This software has a quality that impacts the overall performance of the system.

ISO/IEC 25000 provides a guide to a series of standards that address systems and software quality requirements and evaluation. These standards include the following:

— ISO/IEC 25010, which specifies quality models for systems and software;

— ISO/IEC 25012, which specifies a quality model for data, where the model includes fifteen quality characteristics for data;

— ISO/IEC 25024, which specifies an approach to measuring the quality of data, where the approach makes use of the quality characteristics specified by ISO/IEC 25012.

These standards provide a practical view of data quality in the context of individual systems. This view complements the ISO 8000 series, which addresses wider considerations that include:

— the explicit, formal recognition of the syntactic, semantic and pragmatic types of data quality;

— how organizations can sustain data quality across multiple systems and when data crosses organizational boundaries.

8.5 Standards for industrial data

The ISO 8000 series complements standards that address industrial data. The series enables organizations to create and exploit such data while sustaining the quality of the data.

The following standards specify detailed requirements for industrial data:

— the ISO 10303 series, which specifies requirements for representation, exchange, sharing and archiving of product data;

— the ISO 13584 series, which specifies requirements for representing and exchanging parts libraries;

— the ISO 15926 series, which specifies requirements for integrating the life-cycle data for process plants, including oil and gas production facilities;

— the ISO 22745 series, which specifies requirements for representing open technical dictionaries and applying those dictionaries to master data;

— the ISO/TS 29002 series, which specifies requirements for the exchange of characteristic data.

Each of these standards provides a precise specification of one or more robust elements of a data architecture. Each element is appropriate for one or more particular types of data or one or more particular types of data processing.

EXAMPLE 1 ISO/TS 22745-30 specifies requirements for representing an identification guide that captures a data specification. Each identification guide is an element of the data architecture within an organization. This element enables the organization to receive appropriate data from one or more other organizations.

EXAMPLE 2 By specifying information models for product data, ISO 10303 supports the processing activities of representing, exchanging, sharing and archiving data.

An organization builds an overall data architecture by selecting those elements that meet the specific needs of the organization. This selection process is addressed by the ISO 8000 series.

By focusing on data quality across organizations and when data cross organizational boundaries, the ISO 8000 series establishes an overarching framework. This framework directs the building of effective data architectures and identification of appropriate standards to underpin each architecture.

The ISO 8000 series also specifies:

— requirements that apply to master data and that can be fulfilled by implementing solutions conforming with the ISO 22745 series and the ISO/TS 29002 series (see 7.6);

— guidance that supports organizations intending to apply parts of the ISO 10303 series to improve the quality of product data (see 7.7).

8.6 Standards for types of data other than industrial data

Many different standards specify requirements for data. These standards exist and are necessary to address different combinations of types of data and uses of the data. Each combination determines a generally unique set of requirements on data.

The ISO 8000 series does not address specific requirements that are addressed by other standards. The series instead considers the fundamental nature of data and the consequences of that nature. These considerations are the basis for specifying general approaches to data quality. These approaches support organizations wanting to exploit those other standards while sustaining data quality across all data processing within the organization.

EXAMPLE The ISO 19157 series specifies requirements for data quality of geographic information. ISO 19160-3 specifies requirements for data quality of addresses. ISO/TR 21707 specifies requirements for data quality to support integrated transport information, management and control.

Annex A
(informative)

Document identification

To provide for unambiguous identification of an information object in an open system, the following object identifier is assigned to this document. The meaning of this value is defined in ISO 10303-1.

{ iso standard 8000 part (1) version (2) }

Bibliography

[1] ISO 8000-8, *Data quality — Part 8: Information and data quality: Concepts and measuring*

[2] ISO 8000-51[6], *Data quality — Part 51: Data governance: Exchange of data policy statements*

[3] ISO/TS 8000-60, *Data quality — Part 60: Data quality management: Overview*

[4] ISO 8000-61, *Data quality — Part 61: Data quality management: Process reference model*

[5] ISO 8000-62, *Data quality — Part 62: Data quality management: Organizational process maturity assessment: Application of standards relating to process assessment*

[6] ISO 8000-63, *Data quality — Part 63: Data quality management: Process measurement*

[7] ISO 8000-64, *Data quality — Part 64: Data quality management: Organizational process maturity assessment: Application of the Test Process Improvement Method*

[8] ISO/TS 8000-65, *Data quality — Part 65: Data quality management: Process measurement questionnaire*

[9] ISO 8000-66, *Data quality — Part 66: Data quality management: Assessment indicators for data processing in manufacturing operations*

[10] ISO/TS 8000-81, *Data quality — Part 81: Data quality assessment: Profiling*

[11] ISO/TS 8000-82[7], *Data quality — Part 82: Data quality assessment: Creating data rules*

[12] ISO 8000-100, *Data quality — Part 100: Master data: Exchange of characteristic data: Overview*

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

[14] ISO 8000-115, *Data quality — Part 115: Master data: Exchange of quality identifiers: Syntactic, semantic and resolution requirements*

[15] ISO 8000-116, *Data quality — Part 116: Master data: Exchange of quality identifiers: Application of ISO 8000-115 to authoritative legal entity identifiers*

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

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

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

[19] ISO 8000-150, *Data quality — Part 150: Data quality management: Roles and responsibilities*

[20] ISO/TS 8000-311, *Data quality — Part 311: Guidance for the application of product data quality for shape (PDQ-S)*

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

[22] ISO 9000, *Quality management systems — Fundamentals and vocabulary*

[23] ISO 9001, *Quality management systems — Requirements*

6) Under preparation. Stage at the time of publication: ISO/DIS 8000-51.

7) Under preparation. Stage at the time of publication: ISO/PRF TS 8000-82.

[24] ISO 9004, *Quality management — Quality of an organization — Guidance to achieve sustained success*

[25] ISO 10303 (all parts), *Industrial automation systems and integration — Product data representation and exchange*

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

[27] ISO 10303-21, *Industrial automation systems and integration — Product data representation and exchange — Part 21: Implementation methods: Clear text encoding of the exchange structure*

[28] ISO 10303-59, *Industrial automation systems and integration — Product data representation and exchange — Part 59: Integrated generic resource: Quality of product shape data*

[29] ISO/IEC 10646, *Information technology — Universal coded character set (UCS)*

[30] ISO 13584 (all parts), *Industrial automation systems and integration — Parts library*

[31] ISO 13584-42, *Industrial automation systems and integration — Parts library — Part 42: Description methodology: Methodology for structuring parts families*

[32] ISO 15926 (all parts), *Industrial automation systems and integration — Integration of life cycle data for process plants including oil and gas production facilities*

[33] ISO/TS 15926-4, *Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 4: Initial reference data*

[34] ISO 19157 (all parts), *Geographic information — Data quality*

[35] ISO 19160-3, *Addressing — Part 3: Address data quality*

[36] ISO/TR 21707, *Intelligent transport systems — Integrated transport information, management and control — Data quality in ITS systems*

[37] ISO 22745 (all parts), *Industrial automation systems and integration — Open technical dictionaries and their application to master data*

[38] ISO/TS 22745-10, *Industrial automation systems and integration — Open technical dictionaries and their application to master data — Part 10: Dictionary representation*

[39] ISO 22745-13, *Industrial automation systems and integration — Open technical dictionaries and their application to master data — Part 13: Identification of concepts and terminology*

[40] ISO/TS 22745-30, *Industrial automation systems and integration — Open technical dictionaries and their application to master data — Part 30: Identification guide representation*

[41] ISO 29002[8)], *Industrial automation systems and integration — Exchange of characteristic data*

[42] ISO/TS 29002-5[9)], *Industrial automation systems and integration — Exchange of characteristic data — Part 5: Identification scheme*

[43] ISO/TS 29002-6[10)], *Industrial automation systems and integration — Exchange of characteristic data — Part 6: Concept dictionary terminology reference model*

[44] ISO/TS 29002-10[11)], *Industrial automation systems and integration — Exchange of characteristic data — Part 10: Characteristic data exchange format*

8) Under preparation. Stage at the time of publication: ISO DIS 29002.

9) To be replaced by ISO 29002 (currently at DIS stage).

10) To be replaced by ISO 29002 (currently at DIS stage).

11) To be replaced by ISO 29002 (currently at DIS stage).

[45] ISO 55001:2014, *Asset management — Management systems — Requirements*

[46] ISO/IEC 25000, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guide to SQuaRE*

[47] ISO/IEC 25010, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*

[48] ISO/IEC 25012, *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model*

[49] ISO/IEC 25024, *Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Measurement of data quality*

[50] ISO/IEC 27001, *Information technology — Security techniques — Information security management systems — Requirements*

[51] ISO/IEC 33004, *Information technology — Process assessment — Requirements for process reference, process assessment and maturity models*

[52] IEC 61360-4, *Standard data element types with associated classification scheme for electric components — Part 4: IEC reference collection of standard data element types and component classes*

[53] IEC 62264-1, *Enterprise-control system integration — Part 1: Models and terminology*

[54] GDPR (General Data Protection Regulation), European Parliament and Council of European Union, 2016) Regulation (EU) **2016**/679

[55] Koomen T., Pol M., Test Process Improvement: A step-by-step guide to structured testing, ISBN 0201596245, 1999

(*Continued from second cover*)

The text of ISO standard has been approved as suitable for publication as an Indian Standard without deviations. Certain conventions are, however, not identical to those used in Indian Standards. Attention is particularly drawn to the following:

a) Wherever the words 'International Standard' appear referring to this standard, they should be read as 'Indian Standard'; and

b) Comma (,) has been used as a decimal marker while in Indian Standards, the current practice is to use a point (.) as the decimal marker.

In this adopted standard, reference appears to the following International Standard for which Indian Standard also exists. The corresponding Indian Standard which is to be substituted in its place is listed below along with its degree of equivalence for the edition indicated.

International Standard	*Corresponding Indian Standard*	*Degree of Equivalence*
ISO 8000-2 Data quality — Part 2: Vocabulary	IS/ISO 8000-2 : 2022 Data quality: Part 2 Vocabulary	Identical

Bureau of Indian Standards

BIS is a statutory institution established under the *Bureau of Indian Standards Act*, 2016 to promote harmonious development of the activities of standardization, marking and quality certification of goods and attending to connected matters in the country.

Copyright

Review of Indian Standards

Amendments are issued to standards as the need arises on the basis of comments. Standards are also reviewed periodically; a standard along with amendments is reaffirmed when such review indicates that no changes are needed; if the review indicates that changes are needed, it is taken up for revision. Users of Indian Standards should ascertain that they are in possession of the latest amendments or edition by referring to the website-www.bis.gov.in or www.standardsbis.in.

This Indian Standard has been developed from Doc No.: PGD 18 (22158).

Amendments Issued Since Publication

Amend No.	Date of Issue	Text Affected

BUREAU OF INDIAN STANDARDS

Headquarters:

Manak Bhavan, 9 Bahadur Shah Zafar Marg, New Delhi 110002
Telephones: 2323 0131, 2323 3375, 2323 9402 *Website*: www.bis.gov.in

Regional Offices:		*Telephones*
Central	: 601/A, Konnectus Tower -1, 6th Floor, DMRC Building, Bhavbhuti Marg, New Delhi 110002	2323 7617
Eastern	: 8th Floor, Plot No 7/7 & 7/8, CP Block, Sector V, Salt Lake, Kolkata, West Bengal 700091	2367 0012 2320 9474
Northern	: Plot No. 4-A, Sector 27-B, Madhya Marg, Chandigarh 160019	265 9930
Southern	: C.I.T. Campus, IV Cross Road, Taramani, Chennai 600113	2254 1442 2254 1216
Western	: Plot No. E-9, Road No.-8, MIDC, Andheri (East), Mumbai 400093	2821 8093

Branches : AHMEDABAD. BENGALURU. BHOPAL. BHUBANESHWAR. CHANDIGARH. CHENNAI. COIMBATORE. DEHRADUN. DELHI. FARIDABAD. GHAZIABAD. GUWAHATI. HIMACHAL PRADESH. HUBLI. HYDERABAD. JAIPUR. JAMMU & KASHMIR. JAMSHEDPUR. KOCHI. KOLKATA. LUCKNOW. MADURAI. MUMBAI. NAGPUR. NOIDA. PANIPAT. PATNA. PUNE. RAIPUR. RAJKOT. SURAT. VISAKHAPATNAM.

Published by BIS, New Delhi