SLOVENSKI STANDARD
SIST-TP ISO/TR 8344:2024

01-november-2024

Informatika in dokumentacija - Vprašanja in premisleki za upravljanje zapisov v strukturiranih podatkovnih okoljih

Information and documentation — Issues and considerations for managing records in structured data environments

Information et documentation — Enjeux et considerations pour la gestion des documents d'activité dans les environnements de données structurées

Ta slovenski standard je istoveten z: **ISO/TR 8344:2024**

ICS:

01.140.20	Informacijske vede	Information sciences

SIST-TP ISO/TR 8344:2024 **en**

Technical Report

ISO/TR 8344

Information and documentation — Issues and considerations for managing records in structured data environments

Information et documentation — Enjeux et considerations pour la gestion des documents d'activité dans les environnements de données structurées

First edition
2024-04

Reference number
ISO/TR 8344:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword v

Introduction vi

1 Scope 1

2 Normative reference 1

3 Terms and definitions 1

4 Basic concepts 4
- 4.1 Understanding relationships among data, information, records and knowledge 4
- 4.2 Concept of structured data 5
  - 4.2.1 Physical records and structured data 5
  - 4.2.2 Business systems 5
  - 4.2.3 Databases 5
- 4.3 Different interpretations of the concept of ‘record’ 7
- 4.4 Metadata for records as structured data 7
- 4.5 Characteristics of structured data 7

5 Issues for managing records in structured data environments 8
- 5.1 General 8
- 5.2 Individual records are not self-contained 9
- 5.3 Systems are not designed to capture records 9
- 5.4 Data is often re-used 9
- 5.5 Records controls may not be able to be applied in business systems 9
- 5.6 Records processes may be different in business systems 10
- 5.7 Record meaning may depend on system documentation 10
- 5.8 Ownership and control of records in shared systems may not be clear 10
- 5.9 Disposition of individual records is not feasible 11
- 5.10 Conflict between disposition and the referential integrity rule for relational databases 12
- 5.11 Preserving structured data records over the long-term 12
- 5.12 Data may be highly dynamic 12

6 Concerns about managing records in structured data environments 13
- 6.1 General 13
- 6.2 Approaches to managing records in structured data environments 13
- 6.3 Appraisal of records in structured data environments 13
  - 6.3.1 Appraisal to determine what records need to be captured 13
  - 6.3.2 Appraisal to determine retention requirements 14
- 6.4 Records management by design in structured data environments 14
  - 6.4.1 General 14
  - 6.4.2 Policy approach 15
  - 6.4.3 Technical approach 15
- 6.5 Apply records controls or policies from other applications to business systems 16
- 6.6 Implement a management system for records across business systems 16
  - 6.6.1 Strategic considerations 16
  - 6.6.2 Operational considerations 16

Annex A (informative) **Identification of concepts and characteristics of structured data** 17

Annex B (informative) **Examples of relational database** 18

Annex C (informative) **Concepts of record from management perspective and technology perspective** 23

Annex D (informative) **Template and use cases on records management in structured data environments** 30

Annex E (informative) **Use case-1 as an example** 35

Annex F (informative) **Use case-2 as an example** 40

Annex G (informative) **Use case-3 as an example** **44**
Annex H (informative) **Use case-4 as an example** **49**
Annex I (informative) **By design approach** **53**
Bibliography **60**

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 document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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 46, *Information and documentation*, Subcommittee SC 11, *Archives/records management*.

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

With the digital transformation of government, business, and society, records are increasingly being created in structured data formats in databases, or in business systems that are underpinned by databases. Whilst this has been occurring for several decades, there has been an increase in the volume of data created, stored and analysed with widespread use of sensors and a focus on data driven decision-making. Data structures are also changing, developing from the well-known relational database into new forms which include distributed data systems that are not controlled by a single organization and which may exist across jurisdictions. There is also a significant number of legacy databases that have been decommissioned from active use, but which require ongoing management.

These changes mean that evidence and memory of government, business and society are increasingly represented in structured data formats. This raises issues if structured data is to be trusted as an authoritative source of information, or record, that meets business, legal, and regulatory requirements. As the basis for decision making and operations, structured data becomes the evidence that is subject to e-discovery requirements. If not properly managed, the business, legal, evidential, and information value of structured data can diminish and adversely impact the organization's productivity, compliance, trustworthiness, transparency, accountability and reputation.

Building the capability to manage records in structured data environments has become essential to the governance and management of organizations and communities. There is a growing business need for guidance and recommendations around the design and implementation of adequate policies and procedures to help ensure that records in structured data environments have the attributes of authenticity, reliability, integrity and usability.

Whilst management systems for records as specified in ISO 30301 can be used to ensure that there is appropriate leadership, planning, support, improvement and evaluation with respect to records in structured data environments, there are also specific records control, process and system issues to be considered.

This document provides a landscape review of records management in structured data environments, and identifies issues and considerations for managing records in these environments.

The primary audiences for this document are data policy makers, systems designers, business system owners, data management professionals, database professionals, and the records management professionals working together to ensure the application of appropriate records management approaches, processes, controls and systems in structured data environments.

Information and documentation — Issues and considerations for managing records in structured data environments

1 Scope

This document identifies issues and considerations for managing records in structured data environments.

2 Normative reference

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 30300, *Information and documentation — Records management — Core concepts and vocabulary*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 30300 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/

3.1
attribute
characteristic of an object or entity

[SOURCE: ISO/IEC 2382-36:2019, 3.9.2]

3.2
data
set of characters or symbols to which meaning is or could be assigned

Note 1 to entry: From an ICT perspective, ISO/IEC 2382:2015, 2121272 and ISO 8000-8:2015, 3.1 define data as "reinterpretable representation of information in a formalized manner suitable for communication, interpretation, or processing". In an ICT environment, data is a digital representation of information. It is considered to be the result of how information has been recorded and consists of bits, bytes, characters and pixels.

[SOURCE: ISO 30300:2020, 3.2.4, modified — Note 1 has been added.]

3.3
database
collection of data organized according to a conceptual structure describing the characteristics of these data and the relationships among their corresponding entities, supporting one or more application areas

Note 1 to entry: database: Term and definition standardized by ISO/IEC 2382-1:1993; ISO/IEC 2382-17:1999.

[SOURCE: ISO/IEC 2382:2015, 2121413, modified — Note 2 has been deleted]

3.4
database management system
system, based on hardware and software, for defining, creating, manipulating, controlling, managing, and using databases

Note 1 to entry: The software for using a database may be part of the database management system or may be stand-alone.

Note 2 to entry: database management system; DBMS: term, abbreviation and definition standardized by ISO/IEC 2382-17:1999.

[SOURCE: ISO/IEC 2382:2015, 2121417, modified — Note 3 has been deleted]

3.5
data element
unit of data for which the definition, identification, representation and permissible values are specified by means of a set of attributes

[SOURCE: ISO/IEC 2382-36:2019, 3.8.21]

3.6
entity
any concrete or abstract thing that exists, did exist, or might exist, including associations among these things

EXAMPLE Person, object, event, idea, process, etc.

Note 1 to entry: An entity exists whether data about it are available or not.

[SOURCE: ISO/IEC 2382-36:2019, 3.9.5]

3.7
information
data (3.2) in context with a particular meaning

Note 1 to entry: ISO/IEC 2382:2015, 21212 2 and ISO 8000-8:2015, 3.3 define information as “knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning”.

[SOURCE: ISO 30300:2020, 3.2.7, modified — Note 1 to entry has been added]

3.8
knowledge
maintained, processed and interpreted *information* (3.7)

Note 1 to entry: From ICT and artificial intelligence domain perspectives, ISO/IEC 2382:2015, 2123771 defines knowledge as a “collection of facts, events, beliefs, and rules, organized for systematic use”.

Note 2 to entry: Knowledge is data that is meaningful to particular context.

[SOURCE: ISO 5127:2017, 3.1.1.17, modified — Note 1 to entry and Note 2 to entry have been added.]

3.9
metadata
data about other data, documents, or records <set of data> that describes their content, context, structure, data format, provenance, and/or rights attached to them

Note 1 to entry: See also ISO/TR 14873:2013, 2.29.

[SOURCE: ISO 5127:2017, 3.1.10.26.01]

3.10
metadata for records
structured or semi-structured information, which enables the records processes through time and within and across organizations

[SOURCE: ISO 30300:2020, 3.2.9]

3.11
record
information created or received and maintained as evidence and as an asset by an organization, in pursuit of legal obligations or in the course of conducting business

Note 1 to entry: Records are normally used in plural.

Note 2 to entry: In a management system standard (MSS) implementation, the records created to conduct and direct the management system and to document its implementation are called documented information.

[SOURCE: ISO 30300:2020, 3.2.10]

3.12
records control
instrument for helping in the conduct of records processes

Note 1 to entry: Example of records control include metadata schemas for records, business classification schemes, access and permission rules, and disposition authorities.

[SOURCE: ISO 30300:2020, 3.5.6]

3.13
records management by design
approach in which records management is implemented in the initial design stage and throughout the complete lifecycle of products, processes or services that involve handling record

[SOURCE: Records management by design – Some considerations[47]]

3.14
relational database
database in which the data are organized according to a relational model

Note 1 to entry: relational database: term and definition standardized by ISO/IEC 2382-17:1999.

[SOURCE: ISO/IEC 2382:2015, 17.04.05, modified — Note 2 to entry has been deleted]

3.15
relational database management system
database management system designed for relational databases

Note 1 to entry: In order to use relational data base management systems (RDBMS), it is necessary to represent relational model of data that organizes data (see 4.5) with specific characteristics (tables or relations, unique key, etc.) (see ISO/IEC 25024:2015, Table C.1).

[SOURCE: ISO/IEC 25024:2015, 4.34]

3.16
semi-structured data
aggregate datatype whose components' datatypes and their labels are not predetermined

Note 1 to entry: Semi-structured data are forms of structured data that do not follow structure of data models related to relational databases or other forms of databases.

Note 2 to entry: Examples of semi-structured data include the data that contain HTML tags or other markers to separate semantic elements and to represent hierarchies of records and fields within the data.

[SOURCE: ISO/IEC TS 38505-3:2021, 3.14]

3.17
structured data
data which are organized based on a pre-defined (applicable) set of rules

Note 1 to entry: The predefined set of rules governing the basis on which the data is structured needs to be clearly stated and made known.

Note 2 to entry: A pre-defined data model is often used to govern the structuring of data.

Note 3 to entry: Example of structured data are data contained in relational databases.

[SOURCE: ISO/IEC TS 38505-3:2021, 3.15]

3.18
unstructured data
data which are characterized by not having any structure apart from that record or file level

Note 1 to entry: On the whole unstructured data is not composed of data elements.

EXAMPLE An example of unstructured data is free text.

[SOURCE: ISO/IEC 20546:2019, 3.1.37]

4 Basic concepts

4.1 Understanding relationships among data, information, records and knowledge

The concepts of data, information, records and knowledge are abstract and have different meanings depending on professional perspectives.

From a records management perspective records are information created or received and maintained as evidence and as an asset by an organization, in pursuit of legal obligations or in the course of conducting business.

Records, therefore, are a specific form of information, which require particular management approaches, processes, controls, and systems to ensure they have integrity and provide authentic, reliable and usable evidence.

In the digital environment, records may be in the form of documents or emails, sometimes referred to as files or unstructured data, that are created or communicated as part of business transactions. They are often captured in records systems along with metadata for records.

Records may also be in the form of structured or semi-structured data, captured in business systems that are used to support business processes. Often, these business systems are not designed to capture and manage records. Nevertheless, the organizational need for authoritative evidence of the business processes remains.

Records may form part of the knowledge assets within organizations, especially as documented information.

Figure 1 shows one perspective on the relationship between data, information, and knowledge with respect to meaning. In this perspective, there is an abundance of data which often by itself may not have much meaning. Information then is meaningful data. Meaningful data refers to data which has contributed to achieve purposes or solving tasks. Knowledge is what humans know, understand, and can apply, based on what one has perceived, discovered, and learned from processed, organized, contextualized and meaningful data.

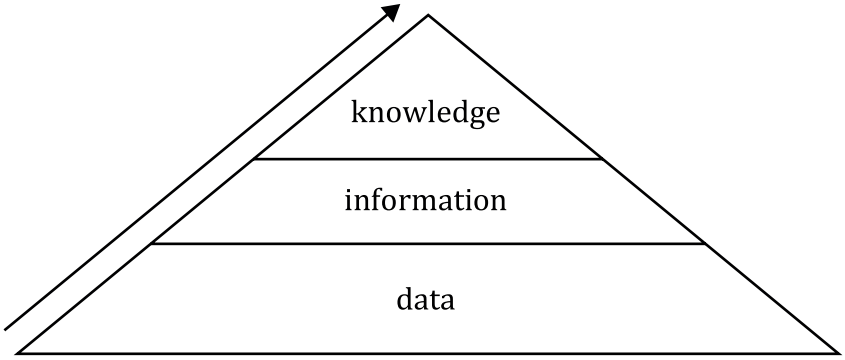

Figure 1 — Relationships of data, information and knowledge

4.2 Concept of structured data

4.2.1 Physical records and structured data

Records managers have always managed records comprised of structured data. In the physical world, these records included:

— registers;

— financial records including ledgers, journals and cash books;

— outputs of instruments, such as seismographs;

— completed forms and charts.

Most of these record forms were replaced by databases as computers were developed to process, organize and record information in digital formats.

4.2.2 Business systems

Initially, data was entered directly into databases. Over time, databases came to underpin business systems where data is entered and processed via a software application and stored in a database. Examples of business systems include human resource management systems, financial systems, contract management systems, case management systems and other transactional systems.

Records managers understand that business systems used to support business purposes often generate data that is needed to serve as evidence of business activity.

However, these systems are usually not specifically designed to manage records.

As outlined in ISO 16175, many business systems generate and store data that can be subject to constant updating (dynamic), are able to be transformed (manipulable), and only contain current data (non-redundant). While business requirements for dynamic, manipulable, and non-redundant data can be entirely legitimate, if records are to serve as reliable evidence of business functions and processes, they need to be fixed and inviolable. Because of the dynamic and manipulable nature of business systems, the capture of records and the ongoing management of their fixity, authenticity, reliability, usability and integrity can be challenging.

4.2.3 Databases

4.2.3.1 General

A database is a collection of data organized according to a conceptual structure describing the characteristics of these data and the relationships among their corresponding entities, supporting one or more application areas.

Databases are typically comprised fields or data items, usually organized in tables.

The content of every field in a row of a table is a data value and conforms to a data type such as string, date, number, etc.

The set of fields or data items treated as a unit is known as a data record or tuple[15].

4.2.3.2 Relational databases

Relational databases have been a common database type for many years. A relational database is a database in which the data are organized according to a relational model.

For example, Bo Wang from Rotterdam owns a house in Gouda, Netherlands, and wants to sell the house using John Johnson's housing broker services.

If the information is stored in a relational database, a data element containing the string 'Gouda' may be the content of a data field of the field type named 'name_town' in a row about the entity Gouda in a table named 'towns'. This table may be part of a relational database named 'houses' owned and managed by John Johnson (or even an association of housing brokers). Such data about towns in a table of a relational database can be the result of a general process of gathering and storing data about towns.

In the same way, the data element 'Bo Wang' can be the content of a data field of the data type named 'name_client' in a row about Bo Wang in a table named 'clients' in another relational database. This kind of data can be the result of John Johnson's specific client intake process. Bo Wang's residential address details are in a table named 'addresses' with a reference from the table 'clients' to this table and another reference from the table 'addresses' to the table 'towns' which includes both 'Gouda' and 'Rotterdam'.

When Bo Wang's house is sold, the sale date and sale price can be added to another table called 'sales'. Thereafter, data already stored in the said tables and databases can be reused in a transaction document prepared by John Johnson and in another transaction document prepared by a notary public.

Annex B provides examples of relational databases for better understanding of their complexity in many and various different types of data models and the predefined set of rules.

4.2.3.3 Master data and transaction data

In some databases, there are two types of data tables: master tables and transaction tables.

Master data is data held by an organization to describe the entities that are both independent and fundamental for the organization and are referenced in transactions. Types of master data include records that describe customers, products, employees, services, etc.

Transaction data is data representing a business transaction. This data may be stored in a transaction table containing data about one type of transaction. Transactions are the result of processes performed which relate to entities recorded in the master data. Transaction data therefore often includes master data that can be used to identify the entities that are party to the transaction.

For example, a credit card transaction relates to entities represented by master data, such as the credit card account at the issuing bank (represented by a credit card number), and the merchant account at the accepting bank (represented by a merchant number)[2].

4.2.3.4 Other forms of databases

The world of data is rapidly changing as familiar relational databases are replaced with alternative data structures, such as graph databases. Graph databases are designed to depict relationships between data points.

The Internet of Things has also resulted in an increase in data being captured by sensors within everyday items. Often, this data is captured in time series databases which record values in simple tables organized by dates and times.