Technical Report

ISO/TR 8344

First edition
2024-04

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

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

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.

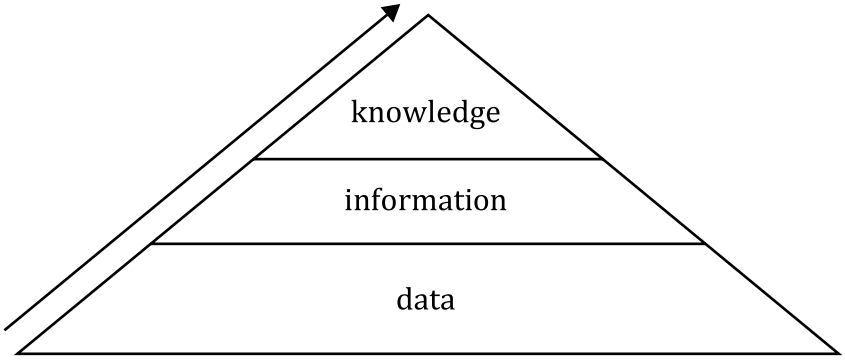

Figure 1 — Relationships of data, information and knowledge

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 the 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.

4.3 Different interpretations of the concept of 'record'

Table C.1 shows there are different interpretations of the concept of "record". For database managers and administrators, a record is a set of related data items treated as a unit[14]. Typically, a data record is a tuple or row of data in a database table.

For records managers, a record comprises all of the information that is created or used in a business transaction, so in a database this might typically comprise related data items from multiple data records across multiple tables.

For example, a person submits an online application by entering information in a web-based form and attaches a scanned document. The form includes fields that require specific items of data, as well as fields that allow free text to be entered. The information is stored as structured, semi-structured and unstructured data. The record comprises the data that was submitted as part of the online application, as well as the metadata about the creation and submission of the application, such as the date and time of creation and the person or agent/actor which created the record.

It may be a complex task to identify which combinations of data elements in a database are considered as a specific record.

Because records managers focus on records as evidence, their aim is to ensure the information content of the record, which may be stored as structured, semi-structured, or unstructured data, can be trusted as authoritative proof of the business transaction.

However, records managers also recognize that records have value as information assets. As information assets, records may be used in subsequent business processes and transactions, generating new records.

4.4 Metadata for records as structured data

In the digital environment, authoritative records are those accompanied by metadata defining their critical characteristics (see ISO 23081-1:2017, Clause 4). According to ISO 15489-1, the metadata of a record itself should be managed as a record. However, it should be noted that the content of a metadata set associated with a given record, unlike the content of the record itself, is not fixed and new metadata are usually added to it over time.

Just like the data itself, the metadata also can be seen and managed both as a set of codes (e.g. when metadata-supporting IT infrastructure such as special databases within electronic document and records management system (EDRMS) systems or communication issues are considered) or as meaningful information that supports integrity, authenticity, usability and confidentiality of records and the operation of the relevant information infrastructure.

There are many kinds of the metadata, of which the "metadata for records" is a subset. In ISO 30300, the metadata for records are defined as "structured or semi-structured information, which enables the records processes through time and within and across organizations". There are numerous other metadata that are not directly related to the records processes (e.g. technical metadata associated with e-mail message and describing its path through the internet).

4.5 Characteristics of structured data

Table A.1 in Annex A identifies characteristics of structured data from representative definitions mapping with ISO 704. There are four characteristics about structured data show in Figure 2:

1) use of a pre-defined set of rules;

2) specified data model;

3) structured in an organized manner or identifiable way;

4) stored in distinct fixed fields.

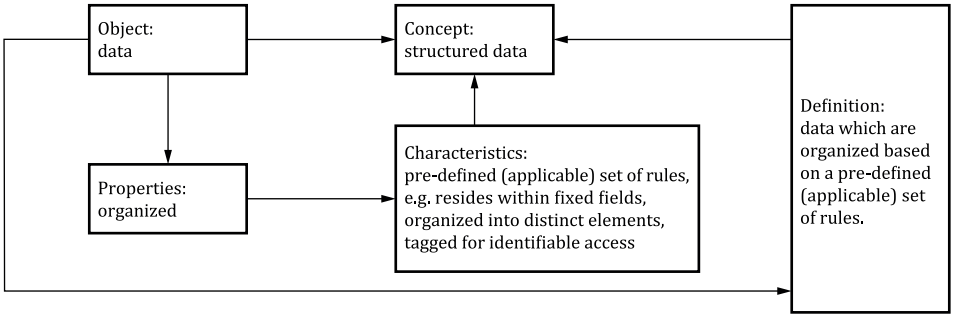

Figure 2 — Formation of structured data concept

Figure 3 illustrates the mapping between the specific characteristics of structured data and the characteristics of records defined in ISO 30300.

The first set of characteristics (e.g. authenticity, usability, reliability and integrity, as defined in ISO 30300) is essential for the categorization of certain data as records.

The second set of characteristics is related to the management of data as records in compliance with the principles and requirements defined in ISO 15489-1 and ISO 30301. These requirements are applicable to all kinds of data (structured, semi-structured, or non-structured) that are managed as records. Besides, managing non-structured data as records need extra work to ensure consistency of data processing and management and the traceability of the management process with adequate metadata for records.

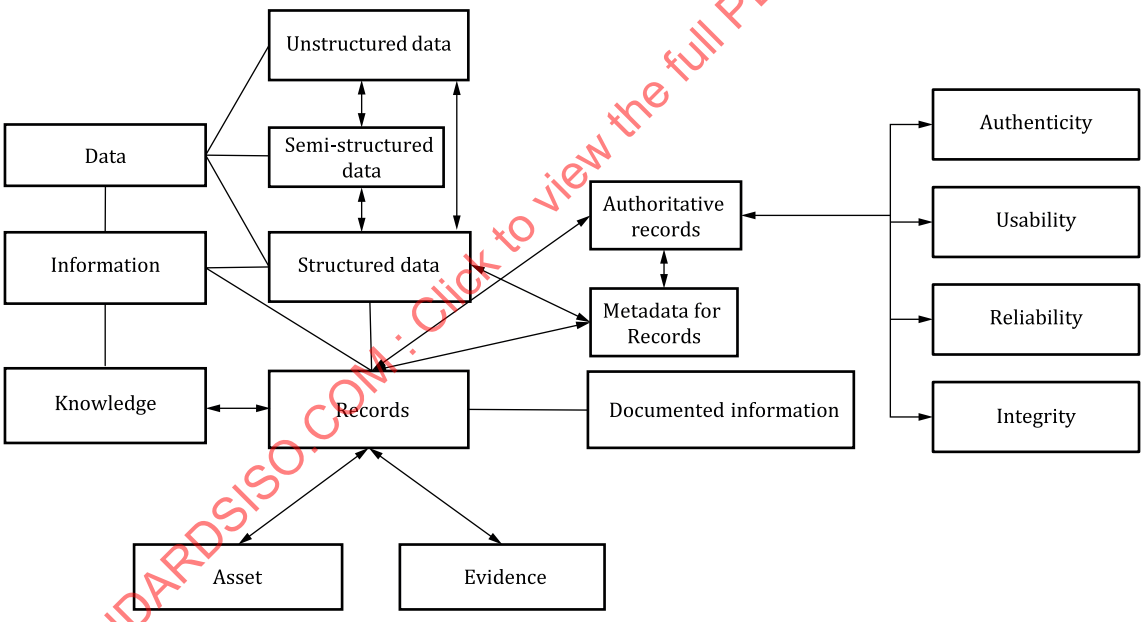

Figure 3 — Mapping concepts of structured data with ISO 30300

5 Issues for managing records in structured data environments

5.1 General

Annex D provides a template for identifying issues and concerns for managing records in structured data environment. The following issues have been validated by four use cases collected in Annex E, Annex F, Annex G, Annex H.

5.2 Individual records are not self-contained

Records in structured data environments are usually part of an overall data structure designed to capture many records of the same type. Generally, all of the records of the same type will be aggregated as a set and cannot be managed individually.

This has implications for the application of records controls and processes as specified in ISO 15489-1.

5.3 Systems are not designed to capture records

In most cases, business systems are still not designed to capture records as records managers hoped by issuing ISO 16175.

Where systems require only current data values, then securing these values against unauthorized changes through authentication and authorization measures may be sufficient.

However, if historical values, such as previous addresses of citizens, have to be available as evidence, it may be challenging to create fixed-content records with data values linked to a point in time.

5.4 Data is often re-used

Once created and stored in a database, data that makes up specific records may be reused in other business processes or transactions, often combined with data from other records.

This re-use may create new records, although the data itself has not changed.

For example, documents or reports generated containing data may need to be saved as separate records, if these documents or reports are used in business processes.

Due to the high level of re-use of data elements, the same data elements can be part of multiple records each of which is subject to numerous retention or disposition requirements.

5.5 Records controls may not be able to be applied in business systems

Records controls may not be able to be applied or implemented in business systems. Nevertheless, some of these can be handled through the management system for records.

Business systems are rarely designed to include functionality that supports business classification schemes and disposition authorities. This might not be an issue where all of the records relate to a single business process and have the same retention or disposition requirements.

Metadata schemas and access and permissions rules are more likely to be able to be incorporated into business systems although the type of metadata and rules are likely to differ compared with records systems, as well as how these controls are implemented.

Metadata for records is different to the typical metadata that is part of structured data environments.

Business systems are unlikely to have all of the metadata for records. For example, business systems may not create explicit identifiers for individual records that are human readable. Also, format or storage metadata may not be explicit. However, this may not be an issue in active records system where records can be retrieved through system functions and the format and storage are known by system users.

The concepts of metadata in ISO 15489-1 seem to be mostly applicable to EDRMS type systems. In business systems, point of capture metadata that documents the circumstances of the creation of the records is typically the name of the agent creating the record and the date and time of record creation.

ISO 15489-1 clarifies that records controls may be designed and implemented in a variety of forms depending on the technological and business environment so these controls can be applied organisation-wide as part of a management system for records, rather than necessarily being applied within records systems.

The management system for records can therefore be used to classify the records in a business system at an aggregate level. Therefore, classification may not occur through application of structured or semi-structured metadata to individual records but could occur through other documentation that shows how the business system relates to business, legal or social contexts.

5.6 Records processes may be different in business systems

Records processes applied to business systems may also be different to those applied in records systems.

In business systems, records are usually captured at the same time as they are created. There is often no significant distinction between creating and capturing records.

ISO 15489-1 seems to assume records are captured into systems that are different from where they are created, which is not the case for many business systems, nor was the case for physical records that preceded them such as financial ledgers.

In distributed ledger technology systems (DLT systems), for example (and like physical ledgers and registers or even diaries), records are created and captured simultaneously with minimal metadata applied to individual records, and no explicit identifier. Metadata was typically applied to the volume containing the individual records.

In business systems, indexing can also occur without explicitly capturing subjects, locations or personal names as indexing metadata. System indexing functions may index the data to enable retrieval of content.

Business systems are typically not designed to support execution of disposition actions, and unlikely to. And yet disposition of records in these systems can still be managed through the management system for records by applying disposition processes and actions at the time of system decommissioning or migration or upgrade.

5.7 Record meaning may depend on system documentation

Records in structured data environments may lose meaning unless codes and additional system documentation is retained, as records may not be meaningful in their own right. Sometimes, metadata for records may or may not serve this purpose.

For example, if an audit entry for a record only includes codes that reference the entities involved in the change event, this entry will be meaningless unless explicit system documentation is retained indicating what the codes referred to.

In addition, it may be useful to retain system documentation showing the data entry screens that were used to capture the records, so that relationships between data items can be understood. All changes to the system can be documented and this documentation retained to help people understand the records.

5.8 Ownership and control of records in shared systems may not be clear

Many business systems are designed to be used by people outside of the organization that established the system, such as systems jointly used by bodies belonging to different levels of government (local, regional, or federal), or systems used by organizations which are part of the same supply chain. Unless agreed up front, there can be a lack of clarity of who owns or is responsible for the records in these shared systems.

Increasingly, people using systems, in which they enter or submit data as part of a business transaction, may also have rights over the data which may be part of an organization's records.

It can be difficult to clearly isolate or separate one organization's records from those of another, each of which may be subject to different legal and regulatory requirements.

Additional challenges are also present in the case of records stored in third-party controlled storage (e.g. cloud) although in general this is not an issue, as usually records are kept separate for joint tenants of shared cloud infrastructure.

Some distributed systems such as blockchain and distributed ledger systems are designed to not necessarily rely on the sole control of one organization, and copies of the ledger are deliberately distributed across the network nodes as part of the system design. This raises additional questions about which are the original records and which are copies.

5.9 Disposition of individual records is not feasible

Technical feasibility of the disposition of individual records is another issue.

Disposition is the range of records processes associated with implementing records retention, destruction or transfer decisions which are documented in disposition authorities or other instruments. Disposition includes both the transfer of records to another organization, permanently archiving records, and the destruction of records at the end of their retention period.

Disposition of records supports effective business, compliance and mitigation of security and privacy risks[12].

For records stored in a relational database, the disposition process requires the removal of specific records from the database.

Specific disposition issues in structured data environments include the following:

— **Identification of the data comprising the records**. Whilst appraisal may have determined the records required as evidence (see 6.3.1), the specific data comprising the records needs to be identified so that the required records to be retained are complete and meaningful. Metadata for records also needs to be identified and retained along with the records.

— **Technical feasibility of destruction.** Individual records in structured data environments cannot generally be disposed of without the risk to the integrity of remaining records in the same database. If the records to be deleted comprise data referring to foreign or secondary keys (see 5.10), this will violate the referential integrity rule for the database. Because individual records are usually part of an overall data structure, it may only be possible to dispose of the complete set of records at the same time. In many cases, the records can only be destroyed during the process of migration of the data or when decommissioning the database, especially when its architecture and design is intended to ensure the immutability of records, e.g. records in a blockchain or distributed ledger system. The disposition process may then be documented outside of the system, in a separate record, rather than being included in metadata for records.

— **Volume of data requiring disposition.** Due to the high volumes of data, manual disposition is not likely to be feasible. This is already problematic in traditional records systems. Policy-driven disposition, implemented by design automation, is even more crucial for structured data environments. Automation can also help in conducting disposition regularly in the course of normal business rather than as ad hoc projects, which is important for the trustworthiness of records management operations and to avoid legal risks in some jurisdictions.

— **Comprehensiveness of disposition.** From both legal and business standpoints, records are typically considered as destroyed or transferred when the controlled record is destroyed or transferred by the organization. However, as with other digital records there may be copies held by the organization (backup copies, paper printouts, etc.).

— **Assurance of the timely, complete and irreversible destruction of records**. In structured data environments that are complex, distributed, shared or controlled by third parties, it can be difficult to get assurance or attestations of the timely, complete, and irreversible destruction of the records. In some cases, this can be mitigated by internal governance and management measures or by inclusion of suitable clauses into contracts or service level agreements with relevant parties.

— **Making information inaccessible is not destruction.** Making information inaccessible by various means (e.g. by destroying corresponding cryptographic keys or cryptographic shredding, or destruction of indexes pointers and links) is not usually considered proper destruction as the information persists and could potentially be discovered with sufficient effort and expenditure. However, in certain situations, this can be the most feasible solution, after legal and information security risks are thoroughly analysed.

If data has to be removed from a records management perspective but cannot be from the database perspective, alternative solutions may be sought such as masking content or shielding content from reading instead of removal. In such situations, one has to accept that such solutions have drawbacks and that sometimes one can only choose an approach of minimizing the risks by accepting the best possible (but not the perfect) solution.

5.10 Conflict between disposition and the referential integrity rule for relational databases

Research has been undertaken into the potential to undertaken disposition in relational databases, and the conclusions identified conflicts with maintaining the referential integrity of databases from the perspective of proper database management[24].

While disposition rules may dictate the deletion of records, referential integrity rules for relational databases may dictate the retention of the same records—potentially creating a conflict.

A relational database satisfies the referential integrity rule for it to be able to operate properly. In a relational database, foreign keys (also referred to as secondary keys) are values in predefined foreign key fields of rows in the tables of the database. These foreign key values refer to values of primary keys of rows in other tables. The primary keys are the identification (ID) codes of rows. The references from secondary keys to primary keys are essential to a relational database's operation. Therefore, each secondary key has to reference a primary key value that exists and a row that exists and contains that primary key value. The deletion of a row referenced by a secondary key damages a relational database, impedes its proper operation, and prevents access to content that are retained. Such a loss of relational database integrity has to be prevented. In practice, it is common that the application controls of a relational database management system (the application for the management of such a database) prevent the execution of commands that would damage the database in this way.

5.11 Preserving structured data records over the long-term

Preservation of document-based digital records (unstructured data) typically relies on conversion of files from proprietary file formats to open file formats or published file formats and there are format registries to assist with ongoing management of files. ISO/TR 22299 provides recommendations for digital file formats for long term storage but this excludes database formats.

There are examples of approaches for converting proprietary databases into standard formats, such as Software Independent Archival of Relational Databases (SIARD), which captures both the data values and the database structure or schema[25].

If data is not converted to non-proprietary formats, licensing policies can prevent the transfer of corresponding software and hardware to archival institutions. Also, archival institutions will be faced with extensive re-engineering efforts in order to ensure access to obsolete formats.

5.12 Data may be highly dynamic

In some technical or scientific environments, there are requirements for continuously changing data values.

Options in these scenarios may typically include:

— sampling the continuously changing values and capturing and storing the resulting sample values;

— determining averages of continuously changing values and capturing and storing the resulting average values;

— storing resulting values in a type of non-relational database that is more suitable for storing ranges of (measurement) values, for example a time-series database.

Recording current output values of smart device sensors in communication networks is a similar challenge.

6 Concerns about managing records in structured data environments

6.1 General

Annex D provides a template for identifying issues and concerns for managing records in structured data environment. The following concerns have been validated by four use cases collected in Annex E, Annex F, Annex G, Annex H.

6.2 Approaches to managing records in structured data environments

Organizations have been creating and managing structured data for decades.

Records managers have largely continued to focus their attention on managing digital documents as records, in records systems designed to support records management processes and apply records controls.

For many years, strategies for managing records have comprised extracting records from databases as data sets or reports and managing them in records systems. This can be a feasible solution in some situations.

However, with the increasing focus on data management, and proliferation of structured data, records managers are increasingly looking to manage records in place within business systems.

For records to have integrity, the information content has to be complete and unaltered. This often contrasts with business systems that require data that is current and regularly updated. Therefore, strategies have to be implemented to ensure the information in the database at the time of the business transaction is available as a record.

Strategies for managing records in place within business systems include:

— design records controls and processes into business systems such as

  — recording the previous value in a field designed to capture historical data and the historical status at a point in time

  — logging changes to data values by using the logging facilities of the database

— apply records controls or policies from other applications to business systems

  — this might involve bringing together the data elements that make up a record to create a distinct digital object that is fixed and unalterable as described in ISO 16175-1

— implement a management system for records across business systems.

Implementation of these strategies can be considered in the design of records management solutions.

These strategies require records managers to work closely with application and database managers and administrators.

Where possible, the requirements for records are incorporated into the design of business systems upfront or when systems are being upgraded. It is more difficult to ensure records will be successfully captured and managed once the system has been built. As such, it is critical for records managers to undertake appraisal.

Once appraisal has been undertaken to determine what records need to be captured and the retention requirements for the records, records managers need to collaborate with system and database designers during the design stage.

6.3 Appraisal of records in structured data environments

6.3.1 Appraisal to determine what records need to be captured

In most cases, records will be required to be captured and retained whenever structured data is used to support an organization's business. The structured data may be needed as evidence for business purposes

or for compliance with legal, regulatory, and other requirements. Structured data records can be subject to various business, legal, regulatory, and societal retention and disposition requirements of one or more jurisdictions.

Appraisal is therefore fundamental for managing records in structured data environments. Strategic analysis of the business context, and assessment of risk, can be undertaken to identify what records need to be created or captured not just to identify retention requirements.

Appraisal can be used to determine which information was received or created as a result of a business process or transaction and is required to be kept and maintained as evidence of that process. From a records management perspective, the record may comprise structured, semi-structured and unstructured data that all relate to the business process. Not all data entered may need to be kept as evidence.

Therefore, appraisal can be used to inform the design of business systems. For business systems that already exist, appraisal can inform solutions for managing records through design, implementation, policy or other tactics.

6.3.2 Appraisal to determine retention requirements

With respect to identifying retention requirements for records in structured data environments, appraisal will be more or less the same as for other records. However, there are several things worth considering.

Firstly, it is not likely to be feasible to specify different retention requirements for individual records, as individual records are part of a larger database comprising many records.

The focus of appraisal will need to be on aggregations of similar types of records. This is no different to appraising a set of physical registers or ledgers which often comprise many individual records maintained together as an aggregation.

Secondly, the current interest in data for decision-making indicates there is likely to be re-use of the data. A data value may have been recorded as part of an initial transaction and may be needed as part of the record as evidence of that transaction. But the same piece of data may then be used for various other purposes to inform decision-making.

In records management, the use of a record in a subsequent transaction typically generates a new record. In this case, each re-use of the data for different purposes can be appraised separately.

However, if data is re-used but not in a business transaction then this type of re-use or potential re-use can inform the retention requirements.

6.4 Records management by design in structured data environments

6.4.1 General

Table I.1 in Annex I identifies objects, characteristics, and application scenarios of 'by design' from definitions of representative sources. Figure I.2 in Annex I shows the general concerns about records management in structured data environment by design in terms of its objective, object, way, quality and application scenarios relevant to records process, records control and information system.

Annex I provides following context for understanding records management by design in structure data environment:

— characteristics of by design mapping with ISO 30300,

— by design as approach and means for information management,

— its process steps for records management,

— concerns about records management by design for structured data,

— multiple axes or dimensions of integrality,

— records management functionality of information systems,

— out-of-box thinking and general preconditions for a successfully design approach.

Based on the analysis of different interpretations of record from management perspective and technology perspective in Annex C and analysis of four uses cases in Annex E, Annex F, Annex G, Annex I, policy approach and technology approach are synthesized.

6.4.2 Policy approach

An approach to records management by design is to define policy for records management.

Records management principles and requirements are identified and specified at the initial design stage and throughout the lifecycle of the business system.

Any initiative to change something in the system also provides an opportunity to include requirements for records management.

High-level process and control steps for a policy approach to records management by design includes but is not limited to:

— using appraisal to identify what constitutes an authoritative record, and which records are to be captured and retained, and when they can be disposed of;

— identifying requirements for the classification, indexing, description, access control disposition, and preservation of these records including regulatory requirements of relevant jurisdictions;

— choosing where (in which information systems) to capture and retain these records;

— designing the solution(s) for capturing, retaining and disposing of the selected records, including fundamental system connections;

— building or purchasing and implementing the solution(s) for capturing, retaining and disposing of these records.

6.4.3 Technical approach

Once records requirements have been determined, a by design approach can be used to include records management rules and processes in the system so they operate automatically in the course of business transactions.

A range of process are needed to ensure that the records management rules and processes:

— cover the whole life cycle of the structured data;

— are standards machine applicable, readable and transferable (SMART);

— are machine executable, interpretable, auditable, and adaptable to the applications of data-enabled technologies such as records management in artificial intelligence (AI) environment.

If choosing to store records in the database itself, a technical by design approach may involve modifying the database's data model, for example, by adding rows and fields for historical and context data, or storing the historical transactional data in relational transaction tables.

Rules for records disposition can also be embedded at the beginning of the data model design to enable the machine to automatically understand the rules and undertake automatic processing. These rules need to take into account possible conflicts between disposition requirements and the referential integrity rule for relational databases (described in 5.10). Because the rules are built into the system's architecture and functionality, seeking the consent of data creators, data managers and data users is no longer necessary.

Functions necessary for managing the records in the database will also need to be defined and this will determine the application logic of the database system for their execution.

Records management in structured data environments by design calls for collaboration among business managers, business analysts, records managers, database administrator, and information system developers.

6.5 Apply records controls or policies from other applications to business systems

Another possible strategy for managing records in structured data environments is to centralise records management in one application and apply records controls or policies to other applications, managing records in place. This is also known as federated information governance. It is more commonly used with unstructured data.

6.6 Implement a management system for records across business systems

6.6.1 Strategic considerations

According to ISO 30301 and ISO 30302, the management system for records (including elements such as records policies and records objectives) is established and compatible with the organization's strategic goals, and be embedded in the organizational governance framework.

With organizations adapting to an increasingly digital society through digital transformation, records management and IT professionals need to align their records and data management strategies as part of the organization's top-level digital transformation plan. This can enable the design and implementation of records management requirements and data and information governance into decisions regarding infrastructure, architecture, processes, systems and data models.

6.6.2 Operational considerations

According to ISO 15489-1 and ISO 16175, records processes and controls are ideally embedded in the records system. Whilst records management requirements and principles can be considered in the initial planning and design stages of business systems, such as in the data model and application logic, as well as at key stages throughout the life cycle of these systems, there are likely to be some records management processes and controls that are not feasible to include as functions in the system. Instead, the management system for records may be able to support records management requirements outside of the specific business system.

Examples include

— **Recording codes used in databases**. If there is a need to identify and document codes and other references, these are ensured to be retained along with the records outside of the system. In an EDRMS, this can also be an issue when the metadata values reference an object ID rather than a meaningful entity, and a search of objects by id would be needed to understand the metadata reference. If the object has been deleted it may not be possible to know what the ID referred to.

— **Capture documentation**. Consider retaining system design documentation, screenshots, system administration, database scheme (schema) and user guides to enable data to be understood in context. These can be retained as separate records.

— **Managing disposition at a system level rather than within systems.** Disposition authorities developed that apply to business systems enable the implementation of disposition actions to be undertaken when systems are upgraded, replaced or decommissioned.

Annex A
(informative)

Identification of concepts and characteristics of structured data

Table A.1 — Identification of characteristics of structured data

No.	Sources	Object	Characteristics	Scenario
1	SOURCE: ISO/IEC 38505-3:2021, 3.15	data	pre-defined (applicable) set of rules;	Use pre-defined data model; relational database
2	Source: John Isaza and Tom Reding, Defensible Disposition of Structured Data, Information Governance Solutions, 2015, see https://www.infogovsolutions.com/disposition-structured-data-1/	data	that resides within fixed fields	Fit into pre-defined Data model; relational database
3	[SOURCE: InterPARES Trust Glossary, see https://interparestrust.org/terminology/term/structured%20data]	data	Organized into distinct elements; stored in distinct fields	data model or standard; relational databases, spreadsheets
4	[SOURCE: ITU-T Y.4500.1 (01/2018)]	data	has a structure; organized in a defined manner	a specified information model
5	Franks 2013 (†560 p.36):	data	Organized in a way identifiable.	
6	Gingrich & Morris 2006 (†358 p. 31):	data	stored in fields and rows	in tables of a relational database; accounting and financial data, customer data, personal data

Annex B
(informative)

Examples of relational database

B.1 Examples of ordinary relational databases

A relational database organizes data into tables where each table contains rows (or tuples), columns and cells. Rows include cells that contain data of entities such as data about persons and products. Columns represent the predefined attributes for these entities such as IDs, first names, last names and product names. Every cell of a row contains the value of one attribute of the entity represented by that row.

The data in rows can refer (or relate) to data in other rows by using primary and foreign keys[1)]. Primary keys are the IDs of rows. Secondary keys are pointers to primary keys of other rows.

The main idea behind the concept of a relational database and its design is to avoid data redundancy. Data redundancy is the term for storing the same data more than once in a considered environment. A relational database is designed to store the same data only once in such a database. This has both advantages and drawbacks.

The relationships between data in a relational database are organized by a set of predefined rules. One of those rules is the referential integrity rule. This rule states that for every secondary key there exists a row with a primary key of the same value so that every secondary key refers to something that exists. If this rule is not met, the database loses its referential integrity and can no longer function properly. This reduces the options for deleting data in a relational database if needed to meet records management requirements for the disposition of records. More about that in subclause 5.10 relating to challenges of referential integrity.

Standard Query Language (SQL) is a query language for accessing and manipulating data stored in relational databases. Formally, SQL is not a part of a relational database, but in practice it is important for working with these databases. An alternative name for a relational database is 'SQL database', especially in use when classifying databases.

B.2 Other types of databases

A common classification of databases is one that is based on the structure of a database. On a high level it divides databases into two categories:

— SQL databases for relational databases;

— NoSQL databases for non-relational databases, where NoSQL stands for Not only SQL.

  The NoSQL category is divided into several subtypes or subcategories. Examples of NoSQL database subcategories are:

  — hierarchical databases containing parent-child relationships, like a family tree;

  — network databases: hierarchical databases with non-hierarchical links;

  — key-value databases containing two-column rows within every row a value and a key for that value;

  — document databases or document-oriented databases, for example Extensible Mark-up Language (XML);

  — graph databases with entities as nodes and connections with other nodes if these connections are meaningful;

1) 'Secondary key' is an alternative term for 'foreign key'.

— column databases with flexible columns;

— time-series databases for storing measured values that change over time.

There will not always be sharp boundaries between them. There are overlaps by databases with characteristics from more than one subcategory.

B.3 More insight in types of databases (optional)

The best way to get some more insight into the mentioned types of databases without being or becoming an expert is by looking at examples of their structure visualized by data model schema like in the images Figure B.1, Figure B.2, Figure B.3, Figure B.4, Figure B.5, Figure B.6, Figure B.7 and Figure B.8.

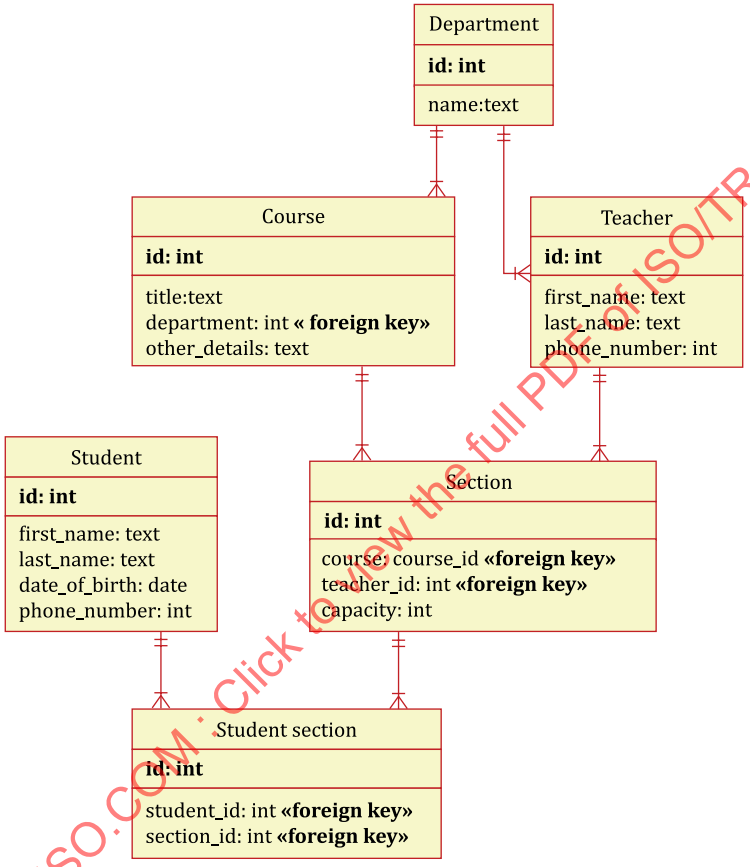

Figure B.1 — Data model of a relational database

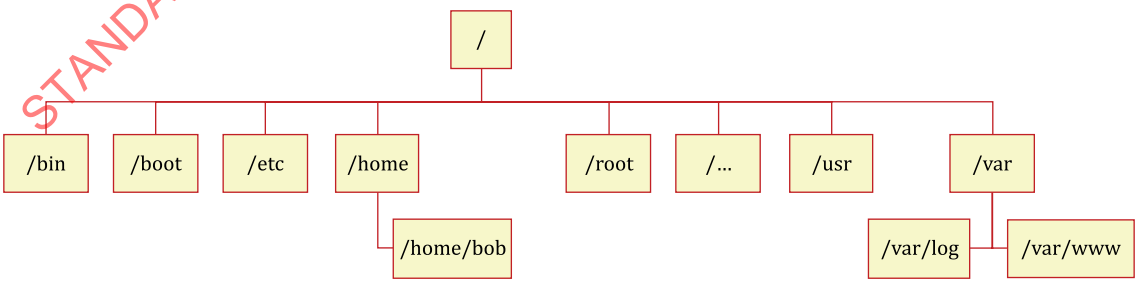

Figure B.2 — Data model of a hierarchical database

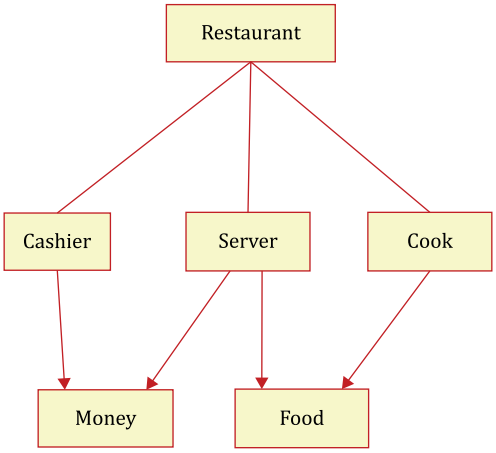

Figure B.3 — Data model of a network database

key:	value
user_id:	f5badc33-5bd7-4b65-a737-65304675f476
color:	blue
repetitions:	3
text:	hello world
data:	{ ... }

Figure B.4 — Data model of a key-value database

ID: breakfast

```
{
        "type": "toast",
        "bread": "whole wheat",
        "spread": [
                "butter",
                "jam"
        ]
}
```

ID: lunch

```
{
        "type": "salad",
        "vegetarian": false,
        "ingredients": [
                "spinach",
                "tomato",
                "cucumber",
                "carrot",
                "dressing": [
                        "olive oil",
                        "vinegar",
                        "honey",
                        "lemon",
                        "salt",
                        "pepper",
                ].
                "tuna",
                "walnuts"
        ].
        "rating": "5 stars",
        "restaurant": "Skylight Diner"
}
```

ID: dinner

```
{
        "type": "pizza",
        "size": "large",
        "toppings": [
                "pepperoni",
                "tomato",
                "sausage"
        ].
        "price": 9.00,
        "presliced": true
}
```

Figure B.5 — Data model of a document database

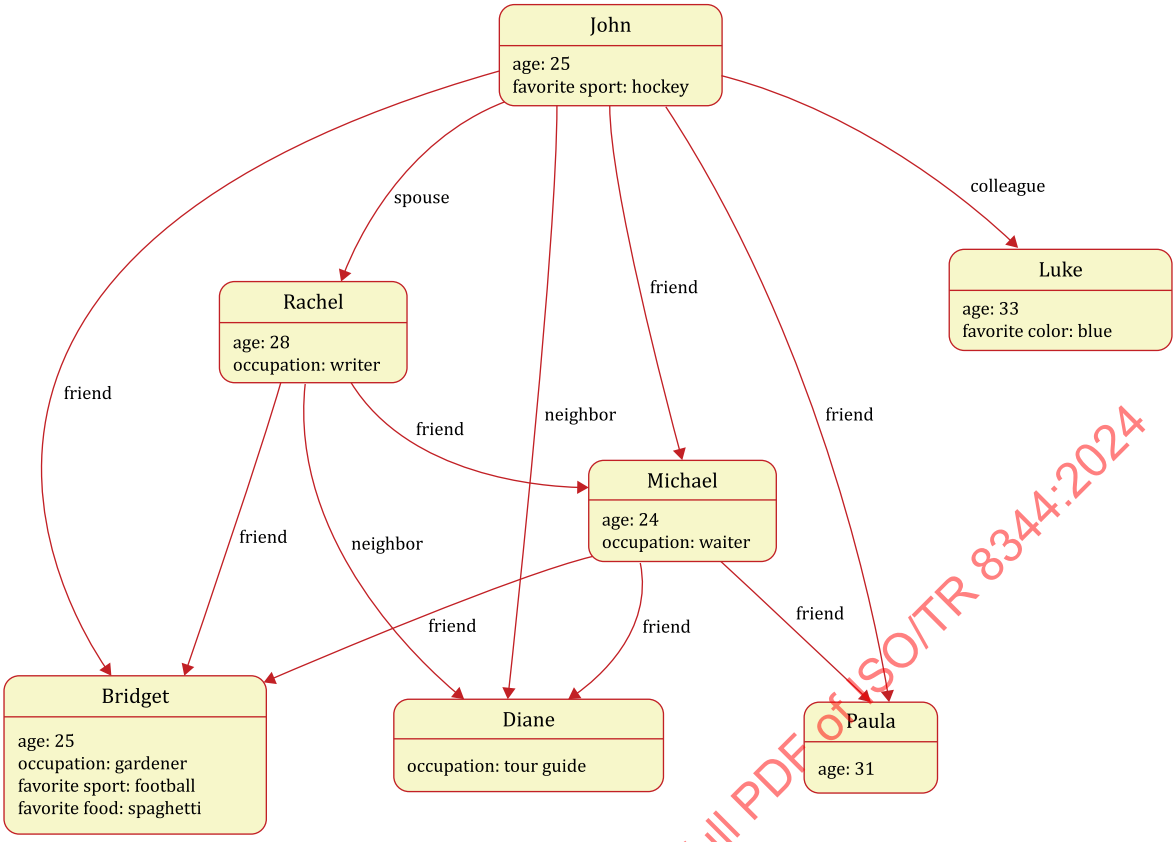

Figure B.6 — Data model of a graph database

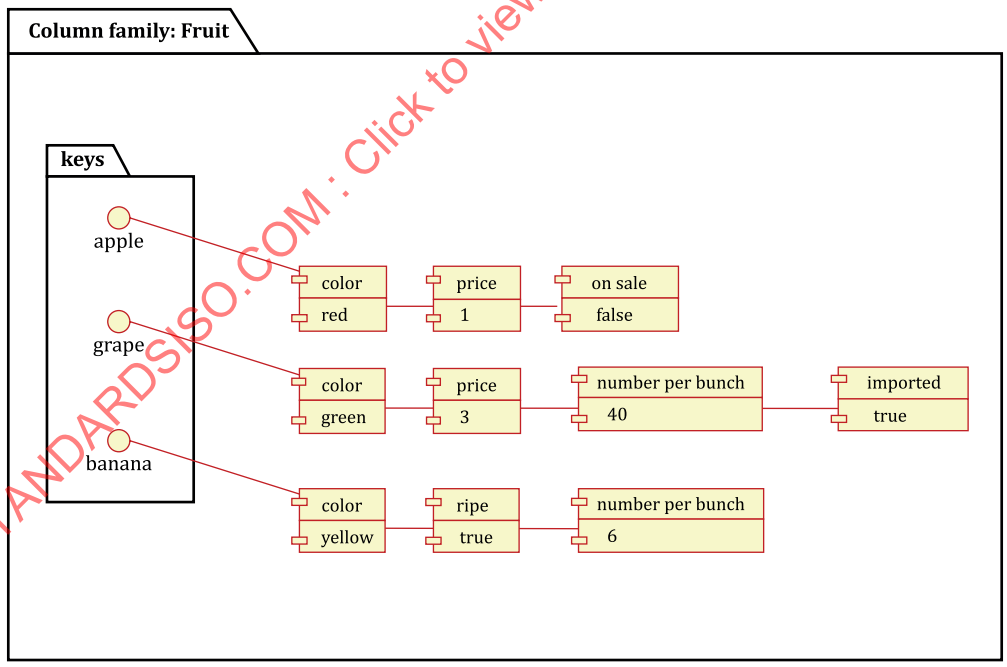

Figure B.7 — Data model of a column database

Time	CPU Temp
2019-10-31T03:48:05+00:00	37
2019-10-31T03:48:10+00:00	42
2019-10-31T03:48:15+00:00	33
2019-10-31T03:48:20+00:00	34
2019-10-31T03:48:25+00:00	40
2019-10-31T03:48:30+00:00	42
2019-10-31T03:48:35+00:00	41

Figure B.8 — Data model of a time-series database

B.4 Some more about relational databases

A relational database is based on the relational model and principles as developed by Edward F. Codd in 1970[46]. In Reference [46], Codd formulated 'twelve rules (0 to 12) for database systems'. Nowadays these rules are considered to be superseded by 'The Third Manifesto'-series of writings on the relational model by Hugh Darwen and C.J. Date[42].

Data (or information) in relational databases are considered as being structured data (or information) according to the definition of the concept 'structured data' as mentioned in 3.17. But note 3 to 3.17 suggests that there are other digital environments that contain structured data too (maybe an XML file according to the specifications in an XML Document Type Definition?).

In the area of database administrators and database managers, rows in relational database tables are also known as 'tuples', 'database records' and in short 'records'. This does not mean that these 'records' are records as defined by TC 46/SC 11 (in ISO 15489-1). In the same area, tables are also named 'relations'.

Annex C
(informative)

Concepts of record from management perspective and technology perspective

Based on the analysis of Table C.1, the conclusion of Figure C.1 was drawn. Figure C.1 shows there are different interpretations of the concept of 'record' from analysis of definitions of record from different standards. There are management perspective and technology perspective on object of record according to different concerns.

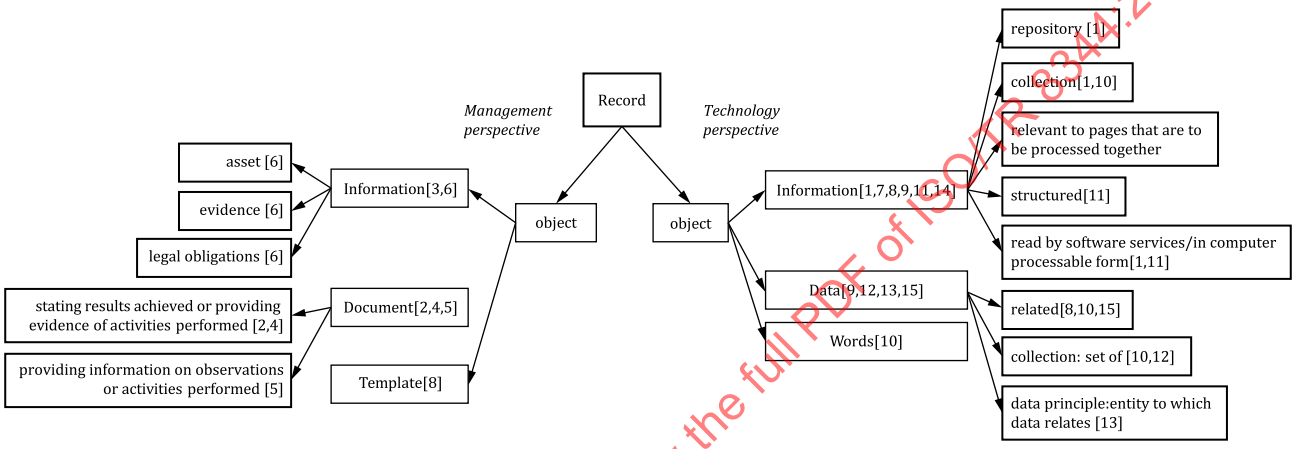

Figure C.1 — Different interpretations of the concept of 'record'

Table C.1 — Concepts of 'record' from different definitions

No.	SOURCE	Definition	Object	Definition or application of the concept of object in its standard or its working group	Characteristics	Perspective
1.	ISO/TS 27790:2009	collection of data	data	**3.17** **data aggregation** process by which information is collected, manipulated and expressed in summary form. Note 1 to entry: Data aggregation is primarily performed for reporting purposes, policy development, health service management, research, statistical analysis and population health studies. **3.18** **data format** arrangement of data in a file or stream **3.19** **data integrity** property that data have not been altered or destroyed in an unauthorized manner **3.20** **data object** collection of data that have a natural grouping and may be identified as a complete entity **3.21** **data structure** manner in which application entities construct the data set information resulting from the use of an information object **3.22** **data subject's consent** any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed **3.23** **data validation** process used to determine if data are accurate, complete or meet specified criteria	collection	Technology
2.	ISO 22263:2008	document stating results achieved or providing evidence of activities performed Note 1 to entry: Adapted from ISO 9000:2005, 3.7.6.	**document**	2.5 **document** information and its supporting medium	stating results achieved or providing evidence of activities performed	Management
3.	ISO/IEC 19794-8:2006	reference and other information about the subject Note 1 to entry: E.g. to access permissions.	reference and other information	ISO/IEC JTC 1/SC 37 Biometrics **biometric information** information needed by the feature extraction and data formatting components of a biometric system to construct the biometric verification or identification data	about the subject	Management

Table C.1 *(continued)*

No.	SOURCE	Definition	Object	Definition or application of the concept of object in its standard or its working group	Characteristics	Perspective
4.	ISO/IEC TR 29110-5-3:2018	document stating results achieved or providing evidence of activities performed	document	3.8 **document** information and the medium on which it is contained EXAMPLE: Record, specification, procedure document, drawing, report, standard. Note 1 to entry: The medium can be paper, magnetic, electronic or optical computer disc, photograph or master sample, or combination thereof. Note 2 to entry: A set of documents, for example specifications and records, is frequently called "documentation". Note 3 to entry: Some requirements (e.g. the requirement to be readable) relate to all types of documents. However, there can be different requirements for specifications (e.g. the requirement to be revision controlled) and for records (e.g. the requirement to be retrievable). [SOURCE: ISO 9000:2015, 3.8.5]	stating results achieved or providing evidence of activities performed	Management
5.	ISO 21043-1:2018	document providing information on observations or activities performed	document	3.6 document Information and the medium on which it is contained Example: specification, procedure document, policy, instruction or form, drawing, record, report, flowchart.	providing information on observations or activities performed	Management
6.	ISO 15489-1:2016	record(s) information created, received and maintained as evidence and as an asset by an organization or person, in pursuit of legal obligations or in the transaction of business	information		as evidence; as an asset; pursuit of legal obligations or in the transaction (3.18) of business	Management
7.	ISO 16612-3:2020	collection of information relevant to pages that are to be processed together Note 1 to entry: A common example of a record is the information about a single recipient for which a collection of pages was generated.	information	ISO/TC 130 Graphic technology **documented information** information required to be controlled and maintained by an organization and the medium on which it is contained Note 1 to entry: Documented information can be in any format and media and from any source. Note 2 to entry: Documented information can refer to the management system, including related processes; information created in order for the organization to operate (documentation); and evidence of results achieved (records).	relevant to pages that are to be processed together	Technology
8.	ISO/IEC 24713-2:2008	template and other information about the end-user EXAMPLE: Access permissions.	template and other information	ISO/IEC JTC 1/SC 37 Biometrics **biometric information** information needed by the feature extraction and data formatting components of a biometric system to construct the biometric verification or identification data (database: structured set of data held in a computer)	about the end-user	Technology

Table C.1 *(continued)*

No.	SOURCE	Definition	Object	Definition or application of the concept of object in its standard or its working group	Characteristics	Perspective
9.	ISO/IEC 18809:2000	Related data treated as a unit of information.	Data; a unit of information	ISO/IEC JTC 1/SC 23 Digitally recorded media for information interchange and storage Data Set: The smallest complete unit of information written to, or received from, the tape.	Related	Technology
10.	ISO/IEC/IEEE 9945:2009	A collection of related data units or words which is treated as a unit.	data units or words	ISO/IEC JTC 1/SC 22 Programming languages, their environments and system software interfaces Data: conforming Prolog cf. conforming Prolog data (source: ISO/IEC 13211-1:1995 Information technology — Programming languages — Prolog — Part 1: General core)	A collection; Treated as a unit.	Technology
11.	ISO 24622-1:2015	structured information that can be read by software services	structured information	ISO/TC 37/SC 4 Language resource management	can be read by software services	Technology
12.	ISO 12615:2004	set of data elements treated as a unit	data	ISO/TC 37/SC 2 Terminology workflow and language coding **data element** **data category** unit of data for which the definition, identification, representation, and permissible values are specified by means of a set of attributes [SOURCE: ISO/IEC 11179-1:2004, definition 3.3.8]	set of; treated as a unit	Technology

Table C.1 *(continued)*

No.	SOURCE	Definition	Object	Definition or application of the concept of object in its standard or its working group	Characteristics	Perspective
13.	ISO/IEC 20889:2018	set of attributes concerning a single data principal	set of data attributes	**3.1** **aggregated data** data representing a group of data principals (3.4), such as a collection of statistical properties of that group **3.2** **attribute** inherent characteristic [SOURCE: ISO 9241-302:2008, 3.4.2] **3.3** **dataset** collection of data [SOURCE: ISO 19115-1:2014, 4.3, modified — The word "identifiable" has been deleted in the definition.] **3.4** **data principal** entity to which data relates Note 1 to entry: The term "data principal" is broader than "PII principal" (or "data subject" as used elsewhere), and is able to denote any entity such as a person, an organization, a device, or a software application. **3.22** **macrodata** dataset comprised of aggregated data **3.23** microdata dataset comprised of records related to individual data principals	concerning a single data principal	Technology
14.	ISO/IEC 11321:1992	Related data treated as a unit of information.	data	ISO/IEC JTC 1/SC 23 Digitally recorded media for information interchange and storage End of Data: The point on the tape at the end of the last valid Data Set recorded on tape. user data: Information that is to be encoded, according to this compression algorithm. [ISO/IEC 22091:2002, 4.30] data byte: An element of user data that is to be encoded.[ISO IEC 22091:2002, 4.4] Processed Data Record(PDR) The data entity resulting from the application of an algorithm to a Logical Data Record.[ISO /IEC 11559:1993, 4.9.1]	Related; treated as a unit of information.	Technology

Table C.1 *(continued)*

No.	SOURCE	Definition	Object	Definition or application of the concept of object in its standard or its working group	Characteristics	Perspective
15.	ISO/IEC 15521:1998	Related data treated as a unit of information.	data	**4.13** **Data Format ID** An identifier specifying which data format is being used on the tape. **4.15** **End of Data** **EOD** The point on the tape at the end of the group which contains the last user data. **4.34** **processed data** A sequence of Codewords which results from the application of processing to data.	Related; treated as a unit of information.	Technology

Annex D
(informative)

Template and use cases on records management in structured data environments

This document gives four use cases about records management in the structured data environment. These use cases which have been collected and analysed were from the Commercial Aircraft Corporation of China, PetroChina, borui Database Company, and State Grid Tianjin Electric Power Company, of which one was from the database manufacturer and three were from the enterprise archives department. Through use case data analysis, it has been found that all the challenges mentioned in Clause 5 were validated, and in addition, some macro management challenges were also identified, such as the lack of file management awareness of the leadership, the risk of technology, the shortage of human resources, the lack of funds, and the constantly updated industry rules, which had brought challenges to records management in the structured data environment.

Table D.1 — Template for collecting and analysis of use case of records management issues and concerns in structured data environment

Row ID	Case content type	Case content	Actions and notes
1.1	Use case code	to be completed by WG22	
1.2	Liaison officer national body → WG22		
1.3	Use case name The use case name needs to reflect the scenario of records management in the structured data environment. e.g. use case and requirements for relational database records scheduling.		
1.4	Additional notes		
About the organization(s)			
2.1	Name of the organization Name of author Country Email		
2.2	Brief description of the organisation's function, key roles, and tasks		
2.3	Additional notes		
About records management of data in structured data environments in general			
3.1	Please provide a brief description of the use case and the associated relational databases. The use case brief introduction needs to reflect the following elements: what is the specific scenario of the structured data environment, what are the associated relational databases (name, type, function, predefined (applicable) set of rules?)? (Less than 200 words)		
3.2	Please provides brief descriptions of types of actors involved in the management of data in relational databases and their roles and the relationships? **Actors who play the roles of managing records in structured data environment in the use case e.g.** (1) Business manager, system developer, data architect, database manager; (2) Data creator (records creation), data processor and controller (records retention, custody and disposition, authority control), data provider (records provider), data user (records user) Please draw a diagram to show relationships between various actors.		
3.3	Does the organization recognize that its relational databases contain data that should be subject to records management rules? If yes, why? Records management rules refer to records management processes and controls in structured data environment. E.g. Capturing, disposing, deleting records in relational database. (what are the rules of ISO 15489-1, ISO 16175, ISO 23081 that are implemented in relational database?) E.g. classification schema, metadata schema, description rule, access and permission rule, disposition authority		
3.4	Is the organization aware of issues and challenges related to applying records management rules to the data in relational databases? If yes, please describe the perceived or identified issues and challenges.		

Table D.1 *(continued)*

Row ID	Case content type	Case content	Actions and notes
	(list of issues identified in WD)		
3.5	Does the organization work with other types of databases than relational databases? If yes: Please provide a brief description of these databases? Are these databases also considered to have challenges in applying records management rules to them? If so, please describe the perceived (or identified) issues and challenges.		
3.6	Additional notes		
About issues and challenges for records management of data in relational databases			
4.1	Issue/challenge 1: **Identifying records in relational databases (appraisal phase).** *Explanatory note:* In a relational database, data elements and combinations of these data elements can be used multiple times for different events and transactions, as well as in multiple information objects about these events and transactions. Examples include requests, decisions, permits, minutes, and reports that serve as evidence of business activities. This results in multiple many-to-many relationships between: — meaningful information elements in the real world; — data elements in the data fields of a relational database. As such, the challenge is to determine which data elements in a relational database are subject to records management rules. Does the organization recognize this as a problem or challenge that applies to the organization? If yes, how does the organization solve this problem or deal with it?		
4.2	Issue/challenge 2: **Conflicting retention periods (appraisal phase).** *Explanatory note:* Using structured data elements in a relational database multiple times in multiple cases and transactions, may result in multiple and conflicting retention periods for the same data elements/element combinations in the database. Does the organization recognize this as a problem or challenge for the organization? If yes, how does the organization solve this problem or deal with it?		
4.3	Issue/challenge 3: **Storing and keeping fixed-content records of current data (capture phase).**		

Table D.1 *(continued)*

Row ID	Case content type	Case content	Actions and notes
	Explanatory note: In a relational database, if structured data elements need to be kept current by updating them, their contents are often not fixed. However, the content of records that serve as reliable evidence of business processes need to be fixed and inviolable. So, it may be challenging to create fixed-content records by capturing values that are initially current and not fixed (see ISO 16175-1:2020, 5.1). Does the organization recognize this as a problem or challenge? If yes, how does the organization solve this problem or deal with it?		
4.4	Issue/challenge 4: **Potential conflict between records management disposition rules for data in a relational database (as records), and the referential integrity rule for a relational database (disposition phase).** *Explanatory note:* If data elements in a relational database are to be considered records and/or if they have been stored as records in such a database, these records will be subject to both: — disposition rules for the records; — the relational databases' referential integrity rule (needed for proper database management and operation of a relational database). As such, a conflict can be created whereby disposition rules dictate the deletion of records, but the referential integrity rule for relational databases dictates their retention. Does the organization recognize this as a problem or challenge? If yes, how does the organization solve this problem or deal with it?		
4.5	Issue/challenge 5: **Collaboration with other professionals on records in a structured data environment.** *Explanatory note:* Record management specialists should collaborate with database managers, database developers, and information architects because of the need to analyse database data models, recognize and solve problems early on, and incorporate records management by design. This can be quite new and challenging if an organization is not yet accustomed to records management by design and a multidisciplinary approach to records management. Does the organization recognize this as a problem or challenge? If yes, how does the organization solve this problem or deal with it?		
4.6	Additional notes		
About more issues and challenges for records management of data in relational databases			
5.1	Has the organization experienced more specific issues/challenges around records management of data in relational databases (other than those above)? If yes, what are they and describe what makes them an issue/ challenge?		
5.2	Do you have any suggestions for WG22 to address other problems or conduct further research?		
5.3	Additional notes		

Table D.1 *(continued)*

Row ID	Case content type	Case content	Actions and notes
In general			
6.1	Additional notes		

Annex E
(informative)

Use case-1 as an example

Table E.1 — Use case of records management issues and concerns in structured data environment: UC-1

Row ID	Case content type	Content
	Date	2023-02-01
	Use case code	UC-1
	Liaison officer WG22 - National body	Xiaomi An (CN)
	Use case name	Managing records in relational databases of telecommunication operator
	Additional notes	
	Name of the organization(s) Name of author Country	Beijing Borrui Data Technology Co., Ltd. Dowson Liu; Zhao Jinghua; Gao Yang; Piao Shenghong China
	Type of organization: role and tasks	Private Corporate Database Company
	Additional notes	
	Brief description of the use case and the relational database or databases involved or in mind for this use case.	Relational databases are widely used in the telecommunication industry, supporting the data analysis of business data. This relational database accesses and processes data directly in memory. No waiting time of loading data from disk to memory is needed. Query processing request is broken up into many smaller tasks, which are distributed intelligently and executed in parallel across nodes for real-time processing and analysis. The distributed architecture of relational database allows a horizontal expansion of hardware. Additional nodes can be added to clusters to meet ever-growing business demand. The unique embedded federated connector system enables users to access various data sources with industry-standard SQL and JDBC interfaces, dispensing with the traditional Extract-Transform-Load (ETL) process and unnecessary data migration.
	Does the organization recognize that data in one or more of its relational databases should be subject to records management rules? If so, why?	Yes, it has been common sense in the industry. The query and analysis of the relevant information, such as account opening information, user call information, and location information for base station positioning are stored in a structured data environment. Organizations have realized that all information is valuable, and some information will be classified as data assets, but there is lack of awareness of records management rules and managing data as records is not adequate.
	Does the organization have a presumption of the issues and challenges it has to face if applying records management rules to the data in its mentioned relational databases? If so, what are these supposed (or discovered) issues and challenges?	The development of industry and technology promotes the improvement of record management awareness, but the considerations of records management rules in the system is largely affected by traditional concepts, capital investment and technical limitations. 1. Leadership awareness: The quality of records management depends on the awareness of records management and the resources input from leadership. The more input, the better the records management. 2. Cost input: Cost control is the most important factor which can affect input to function and performance of system. Only when benefits outweigh costs, records management will be implemented smoothly. 3. Technical risk: Whenever records management are implemented, the new requirements of records management (e.g. immature technology or new technical solution) will bring challenges to the former system technical architecture. 4. Business conflict: The records management may challenges to the original business, retention period rules for example. 5. Other
	Additional notes	

Table E.1 *(continued)*

Row ID	Case content type	Content
	Challenge 1: Identifying records in relational databases (appraisal phase). Explanatory note: The content of data elements and combinations of these data elements in a relational database can be used multiple times for different events and transactions and in multiple information objects about these events and transactions, such as in requests, decisions, permits, minutes and reports which has to serve as evidence of business activities. This leads to multiple n:n relationships between meaningful information elements in the real world and data elements in the data fields of a relational database. In this landscape, the challenge is to determine which data elements in a relational database are subject to records management rules. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Due to lack of identification rules for records, it is hard to identify records from numerous data for databases, which needs records/archives experts, business experts, technology experts, and other relevant departments to cooperate to set identification rules and guide the implementation of records management. 1. Matching relevant data elements with business (records/archives) requirements, and manage access rights, retention periods, and disposal of different data elements in layers. 2. In a structured environment, different management/storage strategies can be set for different types of data, and the system can be flexibly implemented when the business (file) management party puts forward clear requirements at the beginning. 3. The application of core and key data in different scenarios should be considered at the design stage. For example, in order to retain evidence information and avoid business risks, we store user basic account information, call record information location information of base station positioning, and contract information in a structured data environment based on the Archives Law and the Regulations of Telecommunications. The stored data is appraisal as a record according to the following rules: 1. It reflects the result or evidence of a business decision or action. 2. It has legal, financial, management or historical value. 3. It is required by regulations or standards.
	Challenge/issue 2: Conflicting retention periods (appraisal phase). Explanatory note If the content of structured data elements in a relational database is used multiple times in multiple cases and transactions, then this may result in multiple and conflicting retention periods for the same data elements or combinations of data elements in a relational database. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	If the corresponding rules cannot be set at the system design stage, it will be difficult to solve the relevant problems during the process of system operation. Ensure that the system can retain the data according to the business scenario with the consideration of disposition rules and retention schedule at the very beginning of the rules design stage. For example, the call records of users may need to be kept for six months according to the telecommunications regulations, but only for three months according to the local provincial and municipal regulations. In this case, the longest storage period should be followed.
	Challenge/issue 3: Storing and keeping fixed-content records of current data (capture phase). Explanatory note: If structured data elements in a relational database are current and need to be kept current by updating them, their content is not fixed, while the content of records that are to serve as reliable evidence of business processes, need to be fixed and inviolable. So, it may be challenging to create fixed-content records by capturing values that are initially current and not fixed (see ISO 16175-1:2020, 5.1) Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, it is a challenge. The rules should be set in advance and set the capture trigger/ node (including time, frequency, departure event, etc.) If there is a change in this type of data, the system will be triggered to automatically save the current fixed content and store it in the expected format. We use triggers, stored procedures, and other mechanisms to automatically take a snapshot of the data elements when they change and save it in another table as a record. Alternatively, we use specialized record management software or systems that work with relational databases and offer features for creating, capturing, and managing fixed-content records. For example, we have a table in a relational database that stores user information such as name, phone number, ID number, etc. We need to refresh this information periodically to keep it accurate and current. At the same time, we also need to make fixed-content records as reliable proof of our user information management process. We use the database snapshot function to take a snapshot of the database before or after each data update and keep it on the same server instance. This way we query the snapshot to see how the data looked at a certain point in time and compare it with other snapshots.

Table E.1 *(continued)*

Row ID	Case content type	Content
	Challenge/issue 4: Potential conflict between disposition rules from the records management perspective for data in a relational database on the one hand and the referential integrity rule for a relational database on the other (disposition phase). Explanatory note: If data elements in a relational database are to be considered records and/or if these data elements have been stored as records in such a database, these records will be subject to both disposition rules for records and the referential integrity rule for relational databases as needed for proper database management and the operation of a relational database. While disposition rules may dictate the deletion of records, the referential integrity rule for relational databases may dictate the retention of the same records. That is the crux of a potential conflict. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, it is a challenge. The rules of records management shall be incorporated into database processing rules. Once the database processing rules are set, it is difficult to change the rules, which the processing rules of the database must be assured at the design stage. It can meet the requirement of records management by modifying the relational database processing rules. We use the archive function to move a record from the main table to a special area for storing historical data or backup data instead of deleting it from the database when deleting it from the main table. This way, we can retain the original data and reference relationships, and facilitate future recovery or query operations. For example, we use the archive function to handle the relationship between user information and business information. In this case, the user information table and the business information table are linked by a foreign key through the phone number, and when creating the foreign key, specify that deletion is not allowed. In this way, when deleting a record of a canceled user from the user information table, it is not deleted from the database, but moved to a special area for storing historical data or backup data. This way, we retain the original data and reference relationships and facilitate future recovery or query operations.
	Challenge/issue 5: Collaboration with other professionals on records in a structured data environment. Explanation: Record management specialists should collaborate with database managers, database developers and information architects because of the need for analysing database data models, recognizing and solving problems early on and records management by design. This can be quite new and challenging if an organization is not yet accustomed to records management by design and a multidisciplinary approach to records management. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	During the system design process, business experts, records/archives experts, and technical experts including data architects should be fully communicated, design a data management strategy that meets the requirements of records management should be formulated and implemented. However, currently, at the strategic level, both archival management personnel and business personnel lack of guidance of standards for records management to structured data environment. Therefore, it is difficult to formulate relevant rules for records management, which makes it difficult for IT technicians at the operational level to implement relevant requirements for records management. It is necessary to design a practical standard to guide businesses, records/ archives, and system developers to jointly promote records management. We set up a cross-departmental working group or committee to coordinate communication and collaboration among different parties. This ensures that records management experts can timely understand and participate in the process of database design, development, and maintenance, and share records management standards, policies, and best practices with other professionals. We provide training and guidance to enhance the awareness and skills of all parties. We help records management experts and other professionals understand each other's roles, responsibilities, and needs, and improve their ability to manage records in a structured data environment. At the same time, we also use software tools to support the implementation of records management by design. This enables us to leverage technology to automate, simplify and optimize records management processes, such as identifying, classifying, protecting, accessing, and destroying records, and improve data quality and security.
	Additional notes	
	Has your organization experienced more specific challenges/problems in the area of records management of data in relational databases (other than those mentioned here)? If so, what are these challenges/problems and can you explain why they are a challenge/problem?	No

Table E.1 *(continued)*

Row ID	Case content type	Content
	Does the organization work with other types of databases than the relational type? If so, what are these types of databases? And are they also in mind as a challenge when applying records management rules to them? If so, has the organization a presumption of the issues and challenges if doing so? If so, what are these supposed issues and challenges?	Yes, other databases are widely used, such as document databases and graph databases. Document databases mainly provide storage, retrieval, and analysis functions for documents. Compared with relational databases, document databases allow the creation of many different types of unstructured or arbitrary format fields. The main difference from relational databases is that they do not provide support for parameter integrity and distributed transactions, but they are similar to relational databases. They are not mutually exclusive, and they can exchange data with each other, thereby complementing and expanding each other. In this case, the document database is mainly used for storage and query of user credentials, various agreements, contracts, and other information. Graph databases are online database management system that handles create, read, update, and delete operations on graph data models. In this case, the graph database is mainly used to manage the network, access control, and customer relationship. The main challenge faced is the same as relational databases. The main challenge is the lack of awareness of record management rules rather than what types of databases are used.
	Do you have any suggestions for WG22 to address other problems or conduct further research?	1. Due to the lack of awareness of functional requirements for records management during the development of information technology systems, it is hoped that standards for guidance and evaluation will be formulated to guide relevant parties to implement them. 2. Records management in digital environment, including all types of data not just structured data environment, and can not be separately considered from semi-structured or unstructured data environments. 3. Managing records in the SMART area, which are applicable, readable, and transferable by machine, deserves more attention from records management professionals 4. Managing records in AI and in SMART way needs the first priority considerations in future standards development.
	Additional notes	

Annex F
(informative)

Use case-2 as an example

Table F.1 — Use case of records management issues and concerns in structured data enviornment: UC-2

Row ID	Case content type	Content
	Date	2022-12-1
	Use case code	UC-2
	Liaison officer WG22 - national body	Xiaomi An (CN)
	Use case name	Records management of structured data of aircraft products
	Additional notes	
	Name of the organization(s) Name of author Country	COMAC Shanghai Aircraft Design & Research Institute Wang Muliang China
	Type of organization: role and tasks	Manufacturing enterprise Aircraft design
	Additional notes	
	Brief description of the use case and the relational database or databases involved or in mind for this use case.	To accurately and completely record the process of aircraft product development, the firm needs to create a lot of structural product data, such as engineering change requests (ECRs) and engineering change proposals (ECPs). These data are stored in relational databases of about thirty types and are linked to each other, forming indispensable records to prove that the process of aircraft product development has met regulatory requirements and that the product is of acceptable quality. The relational database to generate structural data is a part of the Product data Management (PDM) system. We rely on PDM to create, check, publish, transfer, storage, use and dispose data. The database volume reaches to 50 GB, Queries Per Second (QPS) is 265, and Transactions Per Second (TPS) is 1.15. We face the following challenges: 1. The formulation of data management rules involves multiple departments, which makes it difficult to coordinate among them; 2. Capital investment. Decision-makers may not be able to support this work because of large capital investment and relatively small benefits; 3. Human resources are insufficient. The enterprise may not know how to implement the records management rules due to the lack of professional personnel in the field of records management; 4.Data changes frequently. It may result in the misuse of historical version data. 5.Complex relationship between data. The processing of one type of data may make another type of data unavailable.
	Does the organization recognize that data in one or more of its relational databases should be subject to records management rules? If so, why?	Yes, 1. Implementing the records management rules is conductive to support high-quality information resources for product development, customer service and other businesses, contributing to stability and continuity, and improving the efficiency of operation. 2. To meet the needs of governmental supervision and third-party review, and provide authentic and complete business records. 3. Provide information resources support for the development of new business areas.
	Does the organization have a presumption of the issues and challenges it has to face if applying records management rules to the data in its mentioned relational databases? If so, what are these supposed (or discovered) issues and challenges?	The challenges include: 1. Capital investment. Decision-makers may not be able to support this work because of large capital investment and relatively small benefits; 2. Human resources are insufficient. The enterprise may not know how to implement the record management rules due to the lack of professional personnel in the field of records management;

Table F.1 *(continued)*

Row ID	Case content type	Content
	Additional notes	
	Challenge 1: Identifying records in relational databases (appraisal phase). Explanatory note: The content of data elements and combinations of these data elements in a relational database can be used multiple times for different events and transactions and in multiple information objects about these events and transactions, such as in requests, decisions, permits, minutes and reports which has to serve as evidence of business activities. This leads to multiple n:n relationships between meaningful information elements in the real world and data elements in the data fields of a relational database. In this landscape, the challenge is to determine which data elements in a relational database are subject to records management rules. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes. 1.Determine whether it is within the scope of management according to *the Archive Scope and Retention Period Table*; 2. Hold an appraisal committee meeting to discuss the importance of data and determine whether the records management rules are fully or partially implemented. 3. Formulate relevant institutional document to clarify the record management rules to be followed within the enterprise and their scope of application.
	Challenge/issue 2: Conflicting retention periods (appraisal phase). Explanatory note: If the content of structured data elements in a relational database is used multiple times in multiple cases and transactions, then this may result in multiple and conflicting retention periods for the same data elements or combinations of data elements in a relational database. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, For example, the ECR is a prerequisite for the ECP. The ECP cannot be created without the ECR, the ECR is referenced when the ECP is established, the two belong to the reference relationship, and the relationship is stored in the database. When establishing the retention period, we identified the correlation between the two types of data, so as to set the same storage period.
	Challenge/issue 3: Storing and keeping fixed-content records of current data (capture phase). Explanatory note: If structured data elements in a relational database are current and need to be kept current by updating them, their content is not fixed, while the content of records that are to serve as reliable evidence of business processes, need to be fixed and inviolable. So, it may be challenging to create fixed-content records by capturing values that are initially current and not fixed (source: ISO 16175-1:2020, 5.1) Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, for example, ECP may change after it was published. We take the approach of version control to face this challenge. Each release of a version requires the data of that version to be fixed and at the same time captured and archived.

Table F.1 *(continued)*

Row ID	Case content type	Content
	Challenge/issue 4: Potential conflict between disposition rules from the records management perspective for data in a relational database on the one hand and the referential integrity rule for a relational database on the other (disposition phase). Explanatory note: If data elements in a relational database are to be considered records and/or if these data elements have been stored as records in such a database, these records will be subject to both disposition rules for records and the referential integrity rule for relational databases as needed for proper database management and the operation of a relational database. While disposition rules may dictate the deletion of records, the referential integrity rule for relational databases may dictate the retention of the same records. That is the crux of a potential conflict. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, two rules sometimes conflict. The method we take to deal with it is, When making the deletion decision, the potential impact of the referential integrity rules of the relational database should be fully considered and included in the evaluation scope as the decision input.
	Challenge/issue 5: Collaboration with other professionals on records in a structured data environment. Explanation: Record management specialists should collaborate with database managers, database developers and information architects because of the need for analysing database data models, recognizing and solving problems early on and records management by design. This can be quite new and challenging if an organization is not yet accustomed to records management by design and a multidisciplinary approach to records management. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Yes, collaboration between different professionals is greatly challenging. To cope with the challenges of product data management between different professionals, we established a favorable collaboration mechanism, including: At the beginning of database/software design, seek the opinions of records management experts and incorporate relevant opinions into database/software scheme design and detailed design records; Listen to the opinions of records management experts when dealing with database data; Convene coordination meetings regularly; Incorporate relevant activities into the internal management system.
	Additional notes补充说明	
	Has your organization experienced more specific challenges/problems in the area of records management of data in relational databases (other than those mentioned here)? If so, what are these challenges/problems and can you explain why they are a challenge/problem?	Is the long-term preservation of data records in relational databases dependent on its native database/software or a special electronic records management system? The method adopted is to archive structured data into the electronic records management system with long-term storage function for safekeeping and disposition.
	Does the organization work with other types of databases than the relational type? If so, what are these types of databases? And are they also in mind as a challenge when applying records management rules to them? If so, has the organization a presumption of the issues and challenges if doing so? If so, what are these supposed issues and challenges?	No
	Do you have any suggestions for WG22 to address other problems or conduct further research?	Implementing records management rules in a structured data environment is costly, sometimes even huge. Therefore, the degree of implementing records management rules in structured data environment should be distinguished. Special important data shall be fully implemented; For data of general importance, it can be partially implemented; For unimportant data, it can not be implemented.
	Additional notes	

Annex G
(informative)

Use case-3 as an example

Table G.1 — Use case of records management issues and concerns in structured data enviornment: UC-3

Row ID	Case content type	Content
	Date	2022-12-1
	Use case code	UC-3
	Liaison officer WG22 - national body	Xiaomi An (CN)
	Use case name	Use cases and requirements of relational database in the field of petroleum engineering technology
	Additional notes	
	Name of the organization(s) Name of author Country	China National Petroleum Corporation Jingsheng Fu, Xinghua Su, Zhenghe Song, Guodong Su China
	Type of organization: role and tasks	Chinese central State-owned enterprises; It is a comprehensive international energy company with oil and gas business, engineering technology services, petroleum engineering construction, petroleum equipment manufacturing, financial services, new energy development and other main businesses, and is one of the major oil and gas producers and suppliers in China
	Additional notes	
	Brief description of the use case and the relational database or databases involved or in mind for this use case.	This use case shall be used for data collection, storage and management of relational database in various petroleum engineering technology professional systems. This use case can significantly improve scientific decision-making and economic benefits by storing data generated in production management processes such as automatic collection, manual input, scientific calculation, research results, production management, etc. into relational databases for scientific research, production, management and other applications. The data mainly includes design data of oil drilling, drilling process data, and drilling results data. The relational database used in this use case is Oracle database. Oracle database system is a popular relational database management system in the world at present. The system has good portability, convenient use, and strong functions, and is applicable to various large, medium, small, and microcomputer environments. It is a database scheme with high efficiency, good reliability and high throughput.
	Does the organization recognize that data in one or more of its relational databases should be subject to records management rules? If so, why?	In the petroleum engineering technology industry, there is a strong sense of management of construction achievements data (drilling well history, cementing summary, etc.) as records. However, although the construction process data (caliper data, daily data, etc.) are stored in the structured data environment for query and analysis, the awareness of management of data as records is relatively unconsciousness. However, the petroleum engineering technology industry has realized that all data are valuable and is gradually strengthening the records management requirements and the specification for all data.

Table G.1 *(continued)*

Row ID	Case content type	Content
	Does the organization have a presumption of the issues and challenges it has to face if applying records management rules to the data in its mentioned relational databases? If so, what are these supposed (or discovered) issues and challenges?	With the increasing emphasis on the development of information technology in the petroleum engineering technology industry, some traditional concepts and capital constraints have improved, but still face the impact of industry specifications, technical risks and safety requirements. 1. Industry specifications: due to the possibility that the industry specifications used for some data are outdated or obsolete, the resulting records may face usability and security risks. 2. Technical risk: adding records management to the original old system may cause technical incompatibility or adaptation conflicts and other risks, and will face a lot of compatibility challenges. 3. Security requirements: records management of original data may cause information security risks, such as the record being obtained by unauthorized institutions or individuals. 4. Challenges to traditional concepts and regulations: The traditional thinking and regulations for arching paper records need to be updated.
	Additional notes	
	Challenge 1: Identifying records in relational databases (appraisal phase). Explanatory note: The content of data elements and combinations of these data elements in a relational database can be used multiple times for different events and transactions and in multiple information objects about these events and transactions, such as in requests, decisions, permits, minutes and reports which has to serve as evidence of business activities. This leads to multiple n:n relationships between meaningful information elements in the real world and data elements in the data fields of a relational database. In this landscape, the challenge is to determine which data elements in a relational database are subject to records management rules. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	The petroleum engineering technology industry lacks identification rules for data. At present, through the cooperation of archives, business, technology and other departments, corresponding identification rules have been formulated for some business data according to the actual production and management needs and relevant information systems; The method of rulemaking is mainly based on business logic, but it is difficult to archive process data, such as large amounts of data, high collection frequency, large space consumption, and storage cost performance.
	Challenge/issue 2: Conflicting retention periods (appraisal phase). Explanatory note: If the content of structured data elements in a relational database is used multiple times in multiple cases and transactions, then this may result in multiple and conflicting retention periods for the same data elements or combinations of data elements in a relational database. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	For records with retention period requirements, the data records shall be saved and managed according to the retention period of the business with consideration of the longest retention period in each business system.

Table G.1 *(continued)*

Row ID	Case content type	Content
	Challenge/issue 3: Storing and keeping fixed-content records of current data (capture phase). Explanatory note: If structured data elements in a relational database are current and need to be kept current by updating them, their content is not fixed, while the content of records that are to serve as reliable evidence of business processes, need to be fixed and inviolable. So, it may be challenging to create fixed-content records by capturing values that are initially current and not fixed (see ISO 16175-1:2020, 5.1) Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	At present, the industry captures current data by setting records saving nodes in fixed business processes. For example, we take a certain period of time for well completion as the deadline for archiving, which is to capture and archive records based on the completion of records processing. If such data elements are changed, the new version of the record should be saved again to capture the latest data.
	Challenge/issue 4: Potential conflict between disposition rules from the records management perspective for data in a relational database on the one hand and the referential integrity rule for a relational database on the other (disposition phase). Explanatory note: If data elements in a relational database are to be considered records and/or if these data elements have been stored as records in such a database, these records will be subject to both disposition rules for records and the referential integrity rule for relational databases as needed for proper database management and the operation of a relational database. While disposition rules may dictate the deletion of records, the referential integrity rule for relational databases may dictate the retention of the same records. That is the crux of a potential conflict. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	Modify the length of the relevant field type of the relational database or reassign its content to make it meet designed records management requirements.
	Challenge/issue 5: Collaboration with other professionals on records in a structured data environment. Explanation: Record management specialists should collaborate with database managers, database developers and information architects because of the need for analysing database data models, recognizing and solving problems early on and records management by design. This can be quite new and challenging if an organization is not yet accustomed to records management by design and a multidisciplinary approach to records management. Does the organization recognize this as a problem or challenge? If so, how does the organization solve this problem or deal with it?	In the records management design stage, according to the design requirements, multiple departments such as archives, business and technology are required to jointly discuss and implement. During the design phase, the IT designers form a project team with business and archives related personnel to collaboratively design a strategy that meets the records management requirements of each department. We standardize relevant personnel in the data chain through regulations, such as training for data entry personnel, strengthening the implementation of verification mechanisms for auditors, and participating in the review of major milestone events during the system construction phase.
	Additional notes	
	Has your organization experienced more specific challenges/problems in the area of records management of data in relational databases (other than those mentioned here)? If so, what are these challenges/problems and can you explain why they are a challenge/problem?	None

Table G.1 *(continued)*

Row ID	Case content type	Content
	Does the organization work with other types of databases than the relational type? If so, what are these types of databases? And are they also in mind as a challenge when applying records management rules to them? If so, has the organization a presumption of the issues and challenges if doing so? If so, what are these supposed issues and challenges?	Yes, PI real-time database has been used to store real-time data of production equipment. If records management is needed, data timeliness and space requirements are the main challenges.
	Do you have any suggestions for WG22 to address other problems or conduct further research?	It is hoped that the working group can formulate relevant standards or specifications on records structure in structured data records management to guide the records management design process.
	Additional notes	

Annex H
(informative)

Use case-4 as an example

Table H.1 — Use case of records management issues and concerns in structured data enviornment: UC-4

Row ID	Case content type	Content
	Date	2023-3-8
	Use case code	UC-4
	Liaison officer WG22 - national body	Xiaomi An (CN)
	Use case name	Records Management of Relational Databases in Business Systems
	Additional notes	
	Name of the organization(s) Name of author Country	State Grid Tianjin Electric Power Archives Tingting Liu China
	Type of organization: role and tasks	Central enterprises State Grid Tianjin Electric Power Company is a subsidiary of State Grid Corporation of China. It is responsible for the planning, construction, and operation of the Tianjin power grid, and is committed to providing clean, low-carbon, safe, and efficient electricity and energy supply for the economic and social development of Tianjin.
	Additional notes	
	Brief description of the use case and the relational database or databases involved or in mind for this use case.	The relational database mentioned in this use case exists in the course of various business activities of the company. The company has established corresponding work platforms and relational databases for various businesses, in which relevant functions such as business processing, data query, and information storage are implemented.
	Does the organization recognize that data in one or more of its relational databases should be subject to records management rules? If so, why?	With the approaching of the digital age , the company's various businesses have established a corresponding working platform, on which to implement business handling , data query, information storage and other related functions. But the awareness of considering these data as records to manage and archiving is still relatively weak.
	Does the organization have a presumption of the issues and challenges it has to face if applying records management rules to the data in its mentioned relational databases? If so, what are these supposed (or discovered) issues and challenges?	Yes. There are still various problems and challenges in applying records management to data management in the operation system. It mainly includes the following aspects: 1. Data in partial operation system has strong timeliness and needs to be invoked, processed and analysed at any time. The operating department believes that managing this part of data in accordance with records not only cannot meet the requirements of timeliness but will increase the cost of construction and communication. 2. Data in partial operation system can meet partly requirements of daily work of this business , while managing it as records will face the problem of archiving in the future as well as need to meet the requirements of archive standards, which including the storage format , metadata, electronic signature, encapsulation and backup requirements involved in electronic records management. For professional departments, it is difficult to implement. 3. Data in partial operation system can be preserved for a long time and called at any time. The operating department considers that managing and archiving it as records belongs to repetitive construction, and unfavorable to later utilization.
	Additional notes	