INTERNATIONAL STANDARD

ISO
2709

Fourth edition
2008-07-01

Information and documentation — Format for information exchange

Information et documentation — Format pour l'échange d'information

---

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 2709 was prepared by Technical Committee ISO/TC 46, *Information and documentation*, Subcommittee SC 4, *Technical interoperability*.

This fourth edition cancels and replaces the third edition (ISO 2709:1996), which has been technically revised to incorporate specification of the use of ISO/IEC 10646 using 8-bit Unicode Transformation Format (UTF-8) encoding. Consequently, lengths are specified in terms of octets instead of characters.

Information and documentation — Format for information exchange

1   Scope

This International Standard specifies the requirements for a generalized exchange format which will hold records describing all forms of material capable of bibliographic description as well as other types of records. It does not define the length or the content of individual records and does not assign any meaning to tags, indicators or identifiers, these specifications being the functions of an implementation format.

This International Standard describes a generalized structure, a framework designed specially for communications between data processing systems and not for use as a processing format within systems.

2   Normative references

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

ISO/IEC 646:1991, *Information technology — ISO 7-bit coded character set for information interchange*

ISO/IEC 6429:1992, *Information technology — Control functions for coded character sets*

ISO/IEC 10646:2003, *Information technology — Universal Multiple-Octet Coded Character Set (UCS)*

3   Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1
character
member of a set of elements used for the organization, control, or representation of data

3.2
data field
characters recording the descriptive content of the record

3.3
directory
index to the location of the fields within a record

3.4
directory map
set of parameters specifying the structure of the entries in the directory

3.5
field
variable length portion of the record containing a particular category of data, following the directory and associated with one entry of the directory

NOTE   A field can contain one or more subfields.

3.6
field separator
control character that terminates the directory, record identifier, reference fields and data fields

3.7
tag
three octets associated with a field and used to identify it

3.8
indicator
first data element, if present, associated with a field supplying further information about the contents of the field, about the relationship between the field and other fields in the record, or about the action required in certain data manipulation processes

3.9
octet
group of 8 consecutive bits also referred to as an 8-bit byte; it may represent one character or be part of a representation of a character

3.10
record
collection of fields, including a record label, a directory and data

NOTE   If required, the linking of records and their division into subrecords is allowed and should be implemented as specified in the explicit exchange format.

3.11
record identifier field
characters identifying the record

3.12
record label
section occurring at the beginning of each record providing parameters for the processing of the record

3.13
record separator
control character that terminates the record

3.14
reference field
data that may be required for processing a record

3.15
separating character
control character used to separate and qualify units of data logically, and in some cases hierarchically

3.16
structure
arrangement of the parts constituting a record

3.17
subfield
part of a field containing a defined unit of information

3.18
identifier
data element of one or more characters immediately preceding and identifying a subfield

3.19
subrecord
group of fields within a record which may be treated as an entity

4   Structure of communication format for a record

4.1   General

The general structure of a record is shown schematically in Figure 1. A more detailed structure is shown schematically in Figure 2, which includes four alternatives for the data sections.
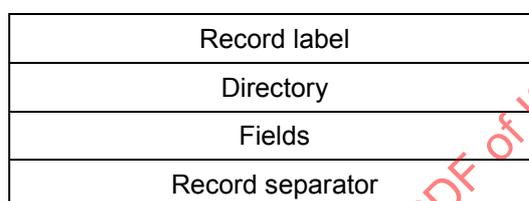
Record label
Directory
Fields
Record separator

Figure 1 — General structure

A record contains the following fixed and variable-length elements in the sequence shown in Figure 2:

a)   record label (fixed length);

b)   directory (variable length);

c)   record identifier field (variable length);

d)   reference fields (variable length);

e)   data fields (variable length);

f)   field separator(s);

g)   record separator.

The directory, record identifier, reference fields and data fields shall each be terminated by a field separator, i.e. separator IS2 in accordance with ISO/IEC 646 or ISO/IEC 6429.

The record shall be terminated by the record separator, i.e. separator IS3 in accordance with ISO/IEC 646 or ISO/IEC 6429.

4.2   Basic character encoding

All data in the record label and directory, and the indicators, subfield identifiers, field separators, and record separators, shall be taken from the character repertoire of ISO/IEC 646, with one octet per character encoding as specified in ISO/IEC 646 or ISO 10646 with UTF-8 encoding.

4.3   Record label

4.3.1   General

The record label shown in Figure 2 shall be fixed in length to 24 octets, each representing one character and the octet positions shall be as defined in 4.3.2 to 4.3.9.

4.3.2   Record length (octets 0 to 4)

The record length shall be defined by the number of octets in the record, including the record label, directory, fields, and the record separator. The length is a five-digit decimal number, right-aligned with zero fill if necessary.

NOTE   The record length described here is a logical record length. For practical reasons relating to machine processing of data, it is sometimes necessary to divide the information into blocks.

4.3.3   Record status (octet 5)

The record status shall be defined by a single character, to be defined in an implementation International Standard, describing the status of a record, for example new or amended. In the absence of an International Standard, special agreement should be reached between the interchange partners.

4.3.4   Implementation codes (octets 6 to 9)

The implementation codes are not defined in this International Standard. Special agreement should be reached between the interchange partners.

4.3.5   Indicator length (octet 10)

The indicator length shall be defined by one decimal digit giving the number of octets of the indicators. If indicators are not used, the indicator length shall be set to zero.

4.3.6   Identifier length (octet 11)

The identifier length shall be defined by one decimal digit giving the number of octets of the identifier. The first or only character of this identifier shall always be IS1 in accordance with ISO/IEC 646 or ISO/IEC 6429. If the identifier is not used, the identifier length shall be set to zero.

4.3.7   Base address of data (octets 12 to 16)

The base address of data shall be defined by five decimal digits, right-aligned with zero fill if necessary, equal to the combined length in characters of the record label and the directory, including the field separator at the end of the directory.

4.3.8   Positions defined by user systems (octets 17 to 19)

Octets positions 17 to 19 shall be defined by user systems.

4.3.9   Directory map

The directory map positions shall be defined by the following.

a)   Octet 20: one decimal digit equal to the length in characters of the length of field part of each entry in the directory.

b)   Octet 21: one decimal digit equal to the length in characters of the starting character position part of each entry in the directory.

c)   Octet 22: one decimal digit equal to the length in characters of the implementation-defined part of each entry in the directory.

d)   Octet 23: reserved for future use.

4.4   Directory

4.4.1   General

The directory shall consist of a variable number of entries, each corresponding to its respective field (record identifier, reference and data fields). The directory shall end with a field separator.

4.4.2   Directory entry

An entry shall consist of the following parts in the given order:

a)   a tag;

b)   the length of the field;

c)   the starting character position;

d)   the implementation-defined part.

The length of the tag shall be three octets. The length in octets of the other three parts in each directory entry shall be given by the directory map (octets 20 to 22 in the record label). All entries in a directory shall have the same structure.

4.4.3   Tag

The tag shall consist of three octets which specify, according to definition in an implementation International Standard, the name of any associated field.

In the absence of an International Standard, special agreement shall be reached between the interchange partners.

4.4.4   Length of field

This length shall be either

a)   the total number of octets [including indicator(s) and field separator] in the field indicated by the preceding tag, or

b)   zero, implying that the directory entry refers to a field whose total length is greater than the largest decimal number (n) which can be stored in the "length" of field part of a directory entry (see Note), or

c)   the number of octets (including field separator) in the final part of a field which has been treated as described in b).

NOTE   In the case of b), the field is regarded as being divided into a number of parts, of which all but the last are of equal length (n). Each part has a corresponding directory entry containing the tag for the field and the starting character position of the part to which the directory entry refers. A length of zero indicates that the directory entry refers to a part of the field which is not the final part and that the length of this part is to be taken as (n).

In the cases described in b) and c), all directory entries which refer to parts of the same field shall be adjacent and in sequence.

4.4.5   Starting character position

The starting character position shall consist of a decimal number giving the position of the first octet of the field identified by the preceding tag, relative to the base address of data [i.e. the starting character position of the first field following the directory is 0 (zero)].